

**This Page Is Inserted by IFW Operations
and is not a part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- **BLACK BORDERS**
- **TEXT CUT OFF AT TOP, BOTTOM OR SIDES**
- **FADED TEXT**
- **ILLEGIBLE TEXT**
- **SKEWED/SLANTED IMAGES**
- **COLORED PHOTOS**
- **BLACK OR VERY BLACK AND WHITE DARK PHOTOS**
- **GRAY SCALE DOCUMENTS**

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

PCTORGANISATION MONDIALE DE LA PROPRIÉTÉ INTELLECTUELLE
Bureau international

DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITE DE COOPERATION EN MATIÈRE DE BREVETS (PCT)

(51) Classification internationale des brevets ⁶ : C12N 15/12, 15/86, 5/10, C07K 14/47, 14/82, A61K 39/395, 48/00, C12Q 1/68, G01N 33/574, 33/68	A2	(11) Numéro de publication internationale: WO 97/22695 (43) Date de publication internationale: 26 juin 1997 (26.06.97)
(21) Numéro de la demande internationale: PCT/FR96/02061 (22) Date de dépôt international: 20 décembre 1996 (20.12.96) (30) Données relatives à la priorité: 95/15146 20 décembre 1995 (20.12.95) FR 96/04853 18 avril 1996 (18.04.96) FR (71) Déposant (pour tous les Etats désignés sauf US): FONDATION JEAN DAUSSET-CEPH [FR/FR]; 27, rue Juliette-Dodu, F-75010 Paris (FR). (72) Inventeurs; et (75) Inventeurs/Déposants (US seulement): TALERMAN, Adam [FR/FR]; 12, rue de la Chaise, F-75007 Paris (FR). AMSON, Robert [FR/FR]; 10, rue Gay-Lussac, F-75005 Paris (FR). COHEN, Daniel [FR/FR]; 3, rue de l'Orme-au-Mesnier, F-91600 Savigny-sur-Orge (FR). (74) Mandataires: MARTIN, Jean-Jacques etc.; Cabinet Regimbeau, 26, avenue Kléber, F-75116 Paris (FR).	(81) Etats désignés: CA, JP, US, brevet européen (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Publiée <i>Sans rapport de recherche internationale, sera republiée dès réception de ce rapport.</i>	
(54) Title: NUCLEOTIDE SEQUENCES, PROTEINS, DRUGS AND DIAGNOSTIC AGENTS FOR TREATING CANCER (54) Titre: SEQUENCES NUCLEOTIDIQUES, PROTEINES, MEDICAMENTS ET AGENTS DIAGNOSTIQUES UTILES DANS LE TRAITEMENT DU CANCER (57) Abstract <p>A nucleotide sequence corresponding to a gene comprising (a) one of sequences SEQ ID 1 to 11, or an equivalent gene which comprises (b) a sequence hybridisable with one of the sequences of (a), (c) a sequence at least 80 % homologous with (a) or (b), or (d) a sequence coding for a protein encoded by a gene according to (a), (b) or (c), or for an equivalent protein, and the use thereof, in particular for controlling cancer as well as for therapeutic follow-up. These genes are in the TSAP (tumor suppressor activated pathway) group, designated TSAP 1 to TSAP 8 and TSAP 3 human (or HUMSIAH) and in TSIP (tumor suppressor inhibited pathway) group, designated TSIP 1 and TSIP 2, both types of genes corresponding to sequences activated or inhibited, respectively, during cellular apoptosis, particularly that induced by p53.</p> (57) Abrégé <p>La présente invention concerne une séquence nucléotidique correspondant à un gène comportant: (a) une séquence selon l'une des IND. SEQ 1 à 11 ou un gène équivalent qui comporte: (b) une séquence s'hybridant avec l'une des séquences selon (a), (c) une séquence présentant au moins 80 % d'homologie avec (a) ou (b), ou (d) une séquence codant pour une protéine codée par un gène selon (a), (b) ou (c) ou pour une protéine équivalente, et leur application notamment dans la suppression du cancer ainsi que dans le suivi thérapeutique. Ces gènes regroupés en TSAP (tumor suppressor activated pathway) et dénommés TSAP 1 à TSAP 8 et TSAP 3 humain (ou HUMSIAH), et en TSIP (tumor suppressor inhibited pathway) et dénommés TSIP 1 et TSIP 2, ces deux types de gènes correspondant respectivement à des séquences induites ou inhibées lors de l'apoptose cellulaire, notamment celles induites par p53.</p>		

UNIQUEMENT A TITRE D'INFORMATION

Codes utilisés pour identifier les Etats parties au PCT, sur les pages de couverture des brochures publiant des demandes internationales en vertu du PCT.

AT	Arménie	GB	Royaume-Uni	MW	Malawi
AT	Autriche	GE	Géorgie	MX	Mexique
AU	Australie	GN	Guinée	NE	Niger
BB	Barbade	GR	Grèce	NL	Pays-Bas
BE	Belgique	HU	Hongrie	NO	Norvège
BF	Burkina Faso	IE	Irlande	NZ	Nouvelle-Zélande
BG	Bulgarie	IT	Italie	PL	Pologne
BJ	Bénin	JP	Japon	PT	Portugal
BR	Brsil	KE	Kenya	RO	Roumanie
BY	Bélarus	KG	Kirghizistan	RU	Fédération de Russie
CA	Canada	KP	République populaire démocratique de Corée	SD	Soudan
CF	République centrafricaine	KR	République de Corée	SE	Suède
CG	Congo	KZ	Kazakhstan	SG	Singapour
CH	Suisse	LJ	Liechtenstein	SI	Slovénie
CI	Côte d'Ivoire	LK	Sri Lanka	SK	Slovaquie
CM	Cameroun	LR	Libéria	SN	Sénégal
CN	Chine	LT	Lituanie	SZ	Swaziland
CS	Tchécoslovaquie	LU	Luxembourg	TD	Tchad
CZ	République tchèque	LV	Lettonie	TG	Togo
DE	Allemagne	MC	Monaco	TJ	Tadjikistan
DK	Danemark	MD	République de Moldova	TT	Trinité-et-Tobago
EE	Estonie	MG	Madagascar	UA	Ukraine
ES	Espagne	ML	Mali	UG	Ouganda
FI	Finlande	MN	Mongolie	US	Etats-Unis d'Amérique
FR	France	MR	Mauritanie	UZ	Ouzbékistan
GA	Gabon			VN	Viet Nam

*SÉQUENCES NUCLÉOTIDIQUES, PROTÉINES, MÉDICAMENTS ET AGENTS
DIAGNOSTICS UTILES DANS LE TRAITEMENT DU CANCER.*

La présente invention concerne la mise en évidence de gènes impliqués dans les voies moléculaires de la suppression tumorale et l'utilisation des gènes ainsi mis en évidence pour le traitement de certains dysfonctionnements géniques, notamment les cancers.

La présente invention a été rendue possible par l'isolement d'ADNc correspondant à des ARN messagers exprimés ou réprimés lors du processus d'apoptose induit par le gène suppresseur p53.

Une analyse globale des événements moléculaires intervenant au cours du cycle cellulaire lors du développement et de l'apoptose cellulaire est nécessaire pour mieux comprendre l'importance du gène p53 dans le processus de suppression tumorale ou, au contraire, de cancérisation.

La transformation d'une cellule normale en cellule tumorale est un processus qui se déroule en plusieurs étapes et qui nécessite une suite d'événements moléculaires. Au niveau physiologique, ces événements se traduiront par une indépendance de la cellule tumorale vis-à-vis des signaux extérieurs ainsi que par une dérégulation interne menant à une croissance incontrôlée.

Deux groupes de gènes sont responsables de cette transformation dite "maligne", d'une part, les oncogènes, d'autre part, les gènes suppresseurs ou anti-oncogènes. Les oncogènes, en raison de leur dérégulation dans le cancer (résultant le plus souvent d'une mutation ou d'une translocalisation) induiront un signal positif qui favorisera la croissance néoplasique. Au contraire, les gènes suppresseurs, du fait de leur délétion, de l'absence de leur expression par mutation du promoteur, par exemple, ou encore de mutations qui modifieront la structure et la fonction de la protéine, seront incapables dans le cancer de fournir le signal qui lui, normalement, devrait freiner cette croissance anormale. En conséquence, le dysfonctionnement des gènes suppresseurs contribue à la transformation néoplasique.

L'objet de la présente invention est l'isolement de gènes ayant normalement une action dans la suppression tumorale et dont il sera alors possible de surveiller et de traiter les éventuels dysfonctionnements.

En particulier, l'isolement de ces gènes permet d'avoir recours à une thérapie génique de remplacement ou bien à la synthèse d'agents pharmacologiques, protéiques ou non protéiques, qui, directement ou

indirectement, par leur action sur les promoteurs, induiront l'activation et l'expression de ces gènes, ou encore la synthèse d'agents pharmacologiques qui permettront de mimer l'effet physiologique de ces gènes suppresseurs.

L'objectif final est, soit d'inhiber la croissance tumorale, ou mieux, d'induire le processus apoptotique de ces cellules tumorales, c'est-à-dire de conduire les cellules tumorales à se "suicider".

La présente invention concerne la mise en évidence de gènes qui sont impliqués dans cette apoptose. En effet, chaque cellule possède en elle un programme de mort physiologique. Il s'agit également d'un processus physiologique qui est impliqué dans le développement afin de maintenir l'homéostasie du corps et de ne pas voir des proliférations cellulaires anormales s'établir, même si, au demeurant, elles n'ont pas de caractère malin.

L'un des gènes suppresseurs les plus importants impliqués dans l'apoptose est le gène p53. Dans sa fonction normale, ce gène contrôle la croissance cellulaire et le processus d'apoptose ; en particulier, c'est ce gène qui bloque la croissance cellulaire et qui doit induire le processus apoptotique afin d'éviter le développement d'un cancer. On a ainsi mis en évidence que des souris nullizygotes pour le p53 étaient beaucoup plus sensibles à la formation de tumeurs. On a également mis en évidence le fait que, dans les cancers, le gène p53 était très souvent altéré et conduisait à la production de protéines incapables de véhiculer le message d'apoptose.

C'est cette particularité qui a été mise en oeuvre dans le cadre de la présente invention.

En effet, la présente invention repose sur la constatation qu'il n'est pas possible, ou du moins qu'il paraît très difficile, de mettre en place une thérapie de substitution directe lors d'un dysfonctionnement du gène p53. En effet, le p53 muté comme il l'est dans le cancer va annuler l'effet physiologique du p53 normal.

Il a donc fallu renoncer, du moins dans un premier temps, à une thérapie de substitution agissant directement au niveau de p53.

La présente invention s'est donc attachée à étudier les gènes situés en aval de p53 afin de "bipasser" la difficulté évoquée précédemment.

Afin d'isoler les gènes activés ou inhibés par le p53 normal (wild-type p53) on a effectué un ratissage global de l'expression des gènes dans une cellule induite en apoptose et dans la même cellule maligne, plus particulièrement dans une cellule exprimant le p53 normal dans sa fonction

et dans une cellule exprimant le p53 muté dont la fonction est oncogénique. La comparaison des gènes exprimés (ARN messagers exprimés dans les deux types de cellule) a permis de mettre en évidence des gènes exprimés différemment, c'est-à-dire exprimés dans l'une des cellules alors qu'ils ne le sont pas dans l'autre (les gènes peuvent être activés ou inhibés).

On en déduit aisément que ces gènes sont impliqués dans le processus de cancérisation, dans un cas par leur absence, et, dans l'autre cas, par leur présence.

Pour cette étude différentielle, la méthode utilisée est la méthode décrite en 1992 par Liang et Pardee (Differential display of eucaryotic mRNA by mean of a polymerase chaine reaction).

Jusqu'à présent, l'isolement des gènes impliqués dans la suppression était effectué soit par clonage positionnel, soit par l'emploi des doubles hybrides. La première méthode permettait, par un calcul statistique, de calculer la plus haute probabilité où pouvait se localiser, au niveau chromosomique, un gène suppresseur candidat pour un type bien particulier de cancer, surtout ceux d'origines familiales. Le système de doubles hybrides permet d'isoler une à une les protéines qui-interagissent avec un gène donné.

L'approche du problème selon la présente invention a permis d'isoler des séquences directement reliées à une fonction. Dès lors, au contraire du séquençage aléatoire des EST, les séquences sont des séquences dont la fonction est connue et qui sont impliquées dans le processus d'apoptose induit par le gène suppresseur p53.

De façon plus précise, cette méthode a été utilisée sur un modèle cellulaire décrit par Moshe Oren, il s'agit de cellules myéloïdes tumorales de souris qui ont été transfectées par un mutant stable du gène p53. En fait, l'expression de ce gène est thermosensible, c'est-à-dire que dans des conditions de culture cellulaire à 37°C la protéine produite est une protéine mutée, c'est-à-dire qu'elle ne peut jouer le rôle de suppresseur de tumeur et donc que la lignée cellulaire correspondante se développe sous forme de cellule maligne, alors qu'à la température de 32°C la protéine p53 exprimée, comme la protéine naturelle, est capable de jouer le rôle de suppresseur et empêche la lignée cellulaire correspondante de devenir maligne.

Cette étude systématique a permis de mettre en évidence les gènes impliqués dans la cascade de suppression induite par p53.

C'est pourquoi la présente invention concerne ces nouvelles séquences et les gènes les comportant ainsi que l'utilisation de ces séquences, tant au niveau du diagnostic qu'au niveau de la thérapie, de même que pour la réalisation de modèles destinés à tester des produits anti-cancéreux.

5 La présente invention concerne tout d'abord une séquence nucléotidique correspondant à un gène comportant :

- (a) une séquence selon l'une des IND.SEQ 1 à 10 ou un gène équivalent qui comporte :
- (b) une séquence s'hybridant avec l'une des séquences selon (a),
- 10 (c) une séquence présentant au moins 80 % d'homologie avec (a) ou (b), ou
- (d) une séquence codant pour une protéine codée par un gène selon (a), (b) ou (c) ou pour une protéine équivalente,

15 et leur application notamment dans la suppression du cancer ainsi que dans le suivi thérapeutique.

De plus, la présente invention concerne un gène humain impliqué dans la cascade de suppression induite par p53 ainsi que l'utilisation des séquences de ce gène, tant au niveau du diagnostic qu'au niveau de la thérapie, de même que pour la réalisation de modèles destinés à tester des produits anti-cancéreux ainsi que leur application à titre d'agent antiviral.

20 La présente invention concerne donc également une séquence nucléotidique correspondant à un gène comportant :

- (a) une séquence selon l'IND.SEQ 11 correspondant au gène TSAP 3 humain ou HUMSIAH (Human Homologue of the Drosophila seven in absentia gene), ou un gène équivalent qui comporte :
- 25 (b) une séquence s'hybridant avec l'une des séquences selon (a),
- (c) une séquence présentant au moins 80 % d'homologie avec (a) ou (b), ou
- (d) une séquence codant pour une protéine codée par un gène selon
- 30 (a), (b) ou (c) ou pour une protéine équivalente,

et leur application notamment dans la suppression du cancer ainsi que dans le suivi thérapeutique.

Concernant les séquences 1 à 11, la présente invention couvre aussi bien la séquence nucléotidique correspondant au gène entier que des fragments de ce gène, notamment lorsqu'ils codent pour une protéine

35 équivalente comme cela sera décrit ci-après.

Les séquences nucléotidiques peuvent être aussi bien de l'ADN que de l'ARN ou des séquences dans lesquelles certains des nucléotides sont non naturels, soit pour améliorer leurs propriétés pharmacologiques, soit pour permettre leur identification.

5 Les séquences mentionnées en (b) (pour les IND.SEQ 1 à 11) sont essentiellement les séquences complémentaires totales ou partielles (notamment pour les cas évoqués précédemment).

Les séquences (a) et (b) (pour les IND.SEQ 1 à 10) permettent non seulement l'accès au gène murin dont elles sont issues, mais également aux
10 gènes humains correspondant par homologie.

Ainsi, l'invention concerne également les séquences nucléotidiques des gènes présentant une forte homologie avec les gènes mentionnés précédemment, de préférence une homologie supérieure à 80 % sur les parties essentielles desdits gènes, soit en général au moins 50 % de la
15 séquence, de préférence l'homologie sera sur ces parties supérieure à 90 %.

Enfin, lorsque lesdits gènes codent pour une protéine, la présente invention concerne également les séquences codant pour la même protéine, compte tenu de la dégénérescence du code génétique, mais également pour les protéines équivalentes, c'est-à-dire produisant les mêmes effets,
20 notamment les protéines délétées et/ou ayant subi des mutations ponctuelles.

Les séquences selon la présente invention sont plus particulièrement les séquences qui sont induites ou inhibées lors de l'apoptose cellulaire, notamment celles induites par p53.

Lesdits gènes sont regroupés en TSAP ou "Tumor Suppressor Activated Pathway" et dénommés de TSAP 1 à TSAP 8 et TSAP 3 humain, correspondant aux IND.SEQ 1 à 8 et 11 (HUMSIAH) respectivement, et en TSIP ou "Tumor Suppressor Inhibited Pathway" et dénommés TSIP 1 et TSIP 2, correspondant aux IND.SEQ 9 et 10.
25

Les caractéristiques des séquences correspondent aux IND.SEQ 1 à
30 10 sont rassemblées dans le tableau ci-annexé.

Les séquences nucléotidiques correspondant aux gènes TSAP (y compris le TSA? 3 humain ou HUMSIAH), sont des séquences exprimées lors du processus d'apoptose alors que lorsqu'ils ne sont pas exprimés le processus d'oncogénèse se poursuit. Il est donc intéressant :

- 35
- de détecter toute anomalie dans le gène correspondant, laquelle peut conduire à une plus grande susceptibilité à l'oncogénèse, et
 - de pouvoir prévoir une thérapie de remplacement.

Il faut d'ailleurs rappeler que ces gènes peuvent intervenir dans d'autres processus que les processus oncogènes ; en effet, p53 est en quelque sorte le gardien de l'intégrité du génome, dans ces conditions les gènes TSAP ou TSIP sont sans doute également impliqués dans cette fonction de contrôle. c'est donc l'ensemble des altérations possibles du génome qui peuvent être redevables de la détection et de la thérapie précédente. Au contraire, les gènes TSIP sont exprimés lors de l'oncogénèse et non lors de l'apoptose, il est donc là aussi intéressant de détecter l'éventuelle anomalie des TSIP et également de prévoir une thérapie d'inhibition/blocage.

La thérapie de remplacement pourra être effectuée par thérapie génique, c'est-à-dire en introduisant le gène TSAP avec les éléments qui permettent son expression in vivo. Les principes de la thérapie génique sont connus. On peut utiliser des vecteurs particuliers, viraux ou non viraux, par exemple des adénovirus, rétrovirus, virus herpès ou poxvirus. La plupart du temps ces vecteurs sont utilisés sous forme défectifs qui serviront de véhicules d'expression de TSAP avec ou sans intégration. Les vecteurs peuvent être également synthétiques, c'est-à-dire mimer des séquences virales, ou bien être constitués par de l'ADN ou de l'ARN nu selon la technique développée notamment par la société VICAL.

Dans la plupart des cas, il faudra prévoir des éléments de ciblage assurant une expression spécifique des tissus ou organes, en effet, il n'est pas possible d'envisager d'activer un phénomène d'apoptose incontrôlé.

La présente invention concerne donc l'ensemble des vecteurs décrits précédemment.

La présente invention concerne également les cellules transformées par un vecteur d'expression tel que décrit précédemment ainsi que la protéine pouvant être obtenue par culture de cellules transformées.

Les systèmes d'expression pour produire des protéines peuvent être aussi bien des systèmes eucaryotes tels que les vecteurs précédents que des systèmes procaryotes dans des cellules de bactéries.

L'un des intérêts de la présente invention est qu'elle a mis en évidence l'implication de plusieurs gènes dans l'apoptose ; ainsi la surexpression de l'un des gènes par thérapie génique peut, pour certains d'entre eux, ne conduire à l'apoptose que les cellules dans lesquelles s'expriment déjà d'autres gènes déréglés, c'est-à-dire des cellules malignes.

La présente invention concerne également, à titre de médicament, un composé assurant l'expression cellulaire d'au moins une des séquences

nucléotidiques précédentes lorsqu'elle est induite lors de l'apoptose cellulaire, notamment des gènes TSAP 1 à TSAP 8 et TSAP 3 humain, ou au contraire assurant l'inhibition de l'expression cellulaire d'au moins une séquence cellulaire telle que décrite précédemment lorsqu'elle est inhibée
5 lors de l'apoptose cellulaire, notamment TSIP 1 et TSIP 2.

Il est, par exemple, possible de prévoir d'autres approches que la thérapie génique, notamment l'utilisation de séquences nucléotidiques en stratégie sens ou antisens, c'est-à-dire pouvant bloquer l'expression de TSIP ou au contraire, agissant en amont, favorisant l'expression de TSAP.

10 On peut également prévoir une stratégie de remplacement directe par apport de protéines correspondant à TSAP ou d'anticorps inhibiteurs correspondant à TSIP.

Enfin, il est possible de prévoir l'utilisation de molécules non protéiques dont l'activité sera d'activer TSAP ou de mimer l'action de son produit d'expression ou bien d'inhiber TSIP ou bien de bloquer l'action de son produit d'expression.
15

Ces produits peuvent être aisément testés sur les cellules modifiées qui sont décrites dans les exemples en introduisant les produits à tester dans la culture cellulaire et en détectant l'apparition du phénomène apoptotique.
20 Dans les stratégies à ADN, ARN ou protéique les produits sont bien entendu élaborés en fonction des séquences qui sont décrites.

La présente invention concerne en particulier l'utilisation des médicaments précédents en tant qu'agent anti-cancéreux.

Mais, le produit du gène TSAP 3 humain (HUMSLI1) est également
25 utile comme agent antiviral, comme cela apparaîtra à la lecture de l'exemple 2. La présente invention concerne donc également l'utilisation des médicaments précédents comme agent antiviral.

La présente invention concerne également à titre d'agent de diagnostic pour la détermination de la prédisposition au cancer, tout ou partie
30 des séquences selon l'invention à utiliser comme sonde nucléotidique ou comme amorce d'amplification, mais également à titre d'agent de diagnostic pour la détermination de la prédisposition au cancer un antigène correspondant à tout ou partie des protéines codées par la séquence selon l'invention ou les anticorps, notamment les anticorps monoclonaux,
35 correspondants, éventuellement après culture.

Les méthodes de diagnostic sont connues, il peut s'agir, par exemple, de techniques de microséquençage des parties variables après

isolement et amplification éventuelle ou des méthodes de détection type RFLP ou d'amplification simple notamment. Les techniques différentielles peuvent, en particulier, permettre de mettre en évidence l'écart entre le TSAP ou TSIP normal et anormal.

5 L'invention concerne également des modèles mettant en oeuvre les séquences précédentes.

Le gène TSAP 3 humain (HUMSIAH) peut être isolé, notamment, en utilisant la méthode PCR ou d'autres méthodes d'amplification en mettant à profit la structure du gène. Il est également possible de synthétiser ce gène
10 par morceau, si nécessaire.

Enfin, l'invention concerne un perfectionnement à la méthode de Liang et Pardee (1) caractérisé en ce que dans l'amplification par PCR on effectue une diminution en palier ("touch down") tel que décrit dans Don et al. (2).

15 D'autres caractéristiques de l'invention apparaîtront à la lecture des exemples ci-après faite notamment en se référant aux figures suivantes :

- Figure 1 - Quantification de l'expression différentielle des ARNm utilisant l'imageur 1200 β . Hybridation aux ARNm dérivés des cellules LTR6 à 37°C et des cellules LTR6 après 4 heures à 32°C. Les nombres en ordonnées de 0 à
20 500 correspondent au comptage détecté par 0,15 mm et sont proportionnels au signal d'hybridation.

C1 : ARNm exprimé également en utilisant un clone sans expression différentielle ;

25 C2 : contrôle positif utilisant la Cycline G et montrant l'induction des ARNm correspondant à 32°C ;

MER-LTR : montre l'induction de cette séquence à 32°C ;

TSAP 1 à TSAP 8 : expression différentielle des 8 ARNm activés dans les 4 premières heures suivant l'induction de l'apoptose ;

30 TSIP 1 et TSIP 2 : expression différentielle des 2 ARNm inhibés dans les 4 premières heures suivant l'induction de l'apoptose.

- Figure 2 - Analyse Northern blot.

A : hybridation avec la sonde TSAP 3 ;

B : hybridation avec la sonde siah 1b de souris ;

35 lignes 1 et 2 : ARNm polyA⁺ de cellules leucémiques myéloïdes M1 (clone S6) cultivées à 37°C et 32°C respectivement ;

lignes 3 et 4 : ARNm polyA⁺ de cellules LTR6 cultivées à 37°C et 32°C respectivement ;

la flèche indique l'expression différentielle du transcrit 1,9 kb de TSAP 3 -
siah 1b de souris ;

panneaux inférieurs : GAPDH ;

C : distribution tissulaire utilisant TSAP 3 comme sonde ;

5 1 : coeur, 2 : cerveau, 3 : rate, 4 : poumon, 5 : foie, 6 : muscle du squelette, 7 :
rein, 8 : testicule ;

les flèches indiquent les transcrits de 1,9 et 2,4 kb ;

panneau inférieur : β -actine.

- Figure 3 - Analyse de l'hybridation in situ avec la sonde TSAP 3 ;

10 A : cellules M1 incubées pendant 4 heures à 32°C et hybridées avec une
sonde antisens TSAP 3 ;

B : cellules LTR6 incubées pendant 4 heures à 32°C et hybridées avec une
sonde sens TSAP 3 ;

15 C : cellules LTR6 incubées à 37°C et hybridées avec une sonde antisens
TSAP 3 ;

D à F : cellules LTR6 cultivées à 32°C pendant respectivement 1, 2 et
4 heures et hybridées à une sonde antisens TSAP 3 ;

la barre dans le panneau A : 10 μ m ;

les flèches indiquent l'accumulation des ARNm TSAP 3 dans le cytoplasme.

20 - Figure 4 - Comparaison entre la séquence d'ADNc de TSAP 1 et la séquence
nucléotidique correspondant à la phospholipase C bêta 4 de rat.

- Figure 5 - Comparaison entre la séquence d'ADNc de TSAP 2 et la séquence
nucléotidique correspondant à la protéine digitée au zinc (ZEN1) localisée
dans le locus Multiple Endocrine Neoplasia (MEN 1).

25 - Figure 6 - Comparaison entre la séquence d'ADNc de TSAP 3 et la séquence
nucléotidique correspondant au gène Drosophila seven in absentia (sina).

- Figure 7 - Comparaison entre le produit des gènes sina de différentes
espèces, humain (HUMSIAH), murin (MMSIAH 1B) et de drosophile
(DROSINA).

30 - Figure 8 - Comparaison entre la séquence d'ADNc de TSIP 2 et la séquence
d'ADNc du transcript S182 murin du gène AD3 impliqué dans la maladie
d'Alzheimer.

MATERIELS ET METHODES

Cultures cellulaires

35 Cellules de leucémie myéloïde M1 (clone S6) et cellules M1
transfectées de façon stable avec un mutant sensible à la température val 135
p53 (LTR6) (3).

Ces cellules sont cultivées sur milieu RPMI 1640 avec 10 % FCS à 5 % de CO₂ à 37°C. Pour la modification de la température, les cultures sont placées dans un second incubateur à 32°C. Pour tous les essais effectués dans cette étude, les cellules sont testées après 12 et 24 heures pour la présence d'apoptose.

Etude des ADNs différentiels

Pour effectuer les tests dans des conditions expérimentales standards et pour obtenir une reproductivité totale des résultats, les modifications suivantes au protocole d'origine (1) ont été effectuées.

On utilise toujours des ARNm polyA⁺ purifiés deux fois sur colonne d'oligodT utilisant Fast Track (Invitrogen, San Diego CA). Après transcription réverse (M-MLV Reverse Transcriptase, Gibco BRL) sur 0.05 µg de polyA⁺ utilisant 20 µM de chacun des dNTP (Boehringer-Mannheim), aucun dNTP additionné n'est ajouté au mélange de PCR final. Un "hot start" à 94°C pendant 5 minutes est effectué avant la PCR (GeneAmp PCR system 9600 Perkin Elmer Cetus). Les échantillons sont refroidis rapidement sur de l'eau glacée. Un "touch down" (2) de 10 cycles de 50°C à 40°C est effectué (94°C 30 secondes - 50°C 1 minute - 72°C 30 secondes), suivi par 35 cycles (94°C 30 secondes - 40°C 1 minute - 72°C 30 secondes) et une extension finale de 5 minutes à 72°C. Les produits de la PCR sont séparés sur gels de polyacrylamide à 6 % non dénaturant (4). Les gels sont exposés sans séchage. Chaque présentation différentielle est effectuée en comparant M156 et LTR6 à 37°C et après 4 heures d'incubation des deux lignées cellulaires à 32°C.

La procédure de présentation différentielle est répétée dans 3 expériences différentes pour confirmer une parfaite reproductibilité.

Les bandes exprimées différentiellement sont découpées à partir du gel, éluées et réamplifiées (1). Les produits de PCR sont sous-clonés en utilisant le système TA-cloning (Invitrogen, San Diego CA) en suivant les indications fournies.

Pour chaque réaction de ligation, 10 clones recombinants sont séquencés en utilisant le système automatique ABI.

Extraction des ARN, analyses et sondes Northern blots

L'ARN total est extrait avec du Trizol (Life Technologies). Les ARN polyA⁺ sont préparés en utilisant le kit OligotexdT (Qiagen, CA). 30 µg de l'ARN total ou 2 µg d'ARN polyA⁺ sont séparés sur agarose 1 %/1 x MOPS / 2 % gel de formaldéhyde, transférés sur membrane de nylon (Hybond N⁺, Appligène, France) comme cela a été décrit précédemment (5). Les Northern blots sont

hybridés avec des sondes marqués au P³² sur les inserts TSAP et TSIP et lavés comme décrit précédemment (5). Pour vérifier l'induction de la fonction du p53 sauvage, les Northern blots sont hybridés avec une sonde cycline G (6). A titre de contrôle pour la quantité d'ARNm chargée, les blots sont
5 hybridés avec une sonde GAPDH. Différents Northern blots (Clontech CA) sont utilisés dans des conditions identiques et hybridés pour le contrôle avec une sonde β -actine. Les produits de RT-PCR pour LTR6 sont amplifiés en utilisant les amorces siah 1b suivantes : 5'CAGTAAACCACTGAAAAACC3' et 5'CAAACCAAACCAAAACCAC3'. Le produit de PCR sous-cloné est utilisé comme
10 sonde contrôle de siah 1b. Les Northern blots sont exposés pendant 10 jours à - 80°C.

Slot blots

La reproductibilité des résultats obtenus par les analyses Northern blot. Les blots sont préparés (Bio-Rad, Hercules CA) en plaçant les produits
15 de PCR (200 ng de Zeta-Probe Blotting Membranes, Bio-Rad, suivant les instructions du fabricant) de clones TSAP et hybridés avec une sonde ADNc marquée au P³² (Superscript II Gibco-BRL, Life Technologies) correspondant à l'ARN des cellules LTR6 incubées à 37°C et ensuite 4 heures à 32°C. Le produit de PCR du clone contenant la cycline G est également déposé sur les
20 membranes et utilisé comme contrôle positif. Les Slot blots sont exposés une nuit à - 80°C.

Analyse quantitative des images

Celle-ci est effectuée en utilisant un imageur 1200 β (Biospace Instruments, Paris, France) sur les deux Northern blots (pour TSIP 1 et TSIP
25 2) et sur les Slot blots pour tous les contrôles ADNsc et TSAP 1 à 8. Pour l'analyse quantitative représentée dans les graphiques de la figure 1 on soustrait un nombre constant de chaque pic. Cette constante est calculée en mesurant la valeur moyenne du bruit de fond dans les slots qui ne contiennent pas d'ADNc. Les résultats du β imageur ont été obtenus en
30 comptant les slot blots une nuit et en les confirmant par autoradiographie avec des temps variables d'exposition. Ces autoradiogrammes montrent les mêmes variations qualitatives relatives entre les activités à 32°C et à 37°C que les mesures effectuées avec le β imageur.

Hybridation in situ (7, 8)

35 Les cellules sont lavées 3 fois dans un tampon phosphate salin (PBS) "cytospinned" et fixées par du paraformaldéhyde à 4 % dans PBS pendant 10 minutes puis conservées dans l'éthanol à 70 %. Des transcrits

d'ARN marqués à la digoxigénine-11-urédine-5'-triphosphate (DIG) et à la biotine-11-UTP de TSAP 3 sont utilisés dans les analyses suivant la procédure décrite précédemment (Boehringer-Mannheim). Pour la détection des souches marquées à la digoxigénine hybridée les tranches sont incubées dans SAD-10 (10 nm d'anticorps anti-DIG de mouton marqués à l'or à 1/1000 de dilution, Biocell UK). L'analyse est effectuée en utilisant de la microscopie à laser confocal.

EXEMPLE 1

L'étude différentielle des ADNc par la méthode de Liang et Pardee permet de disposer d'un outil très puissant et efficace pour détecter les variations dans l'expression des gènes. Néanmoins, il a fallu modifier le protocole original comme cela a été indiqué précédemment afin d'écartier certains problèmes de reproductibilité observés lorsque l'on applique la méthode telle qu'elle est décrite à l'origine.

On a pu mettre en évidence une reproductibilité totale lorsque dans la méthode PCR on introduit un "hot start" suivi par un "touch down".

Les bandes exprimées différentiellement après isolement et réamplification sont néanmoins souvent contaminées par des bandes provenant des ARN qui migrent dans les régions voisines de l'ADNc, si l'on utilise directement ces sondes sur des Northern blots ceci conduit à des erreurs. On a donc sous-cloné les produits de seconde PCR et fait effectuer les analyses des Northern blots utilisés à défaut de recombinant à sonde simple. Le séquençage systématique d'au moins 10 sous-clones recombinants pour chaque bande sélectionnée a montré qu'il était très efficace pour sélectionner les clones d'intérêt.

Le gène p53 est, dans l'état actuel de nos connaissances, le suppresseur tumoral qui est muté dans le plus grand nombre de cancers d'origines très diverses, et l'utilisation du mutant sensible à la température val-135 p53 s'est déjà montrée précédemment fournir des informations très importantes concernant le fonctionnement du p53 sauvage en induisant, soit l'arrêt de la croissance cellulaire en phase G-1, soit l'initiation du programme de mort cellulaire.

Jusqu'à maintenant, les voies moléculaires en amont et en aval de p53 et qui conduisent à la suppression tumorale étaient encore peu claires.

Jusqu'à maintenant un certain nombre de gènes en aval de p53 ont été identifiés, il s'agit notamment de gadd 45, mdm 2, mck, "Mouse endogenous retrovirus" LTR, p21-waf et Cycline G.

La présente invention a permis de mettre en évidence l'existence de 11 gènes qui sont exprimés différenciellement dans les cellules exprimant le p53 sous sa forme suppresseur actif ou bien dans des cellules tumorales exprimant le gène p53 non actif.

5 La figure 1 montre la quantification des signaux d'hybridation correspondant à l'expression différentielle de 8 de ces gènes qui sont activés à 32°C, c'est-à-dire dans lesquels la fonction de p53 sauvage est activée et conduit donc à l'apoptose des cellules, ces gènes qui sont activés seront
10 dénommés ci-après TSAP (pour Tumor Suppressor Activated Pathway), par contre on constate que dans deux expériences 2 gènes exprimés à 37°C sont en partie inhibés à 32°C, ce qui impliquerait qu'ils sont inhibés durant la mort cellulaire programmée, ces gènes ont été dénommés TSIP (pour Tumor Suppressor Inhibited Pathway).

L'analyse des homologies des différentes séquences activées de
15 TSAP 1 à TSAP 3 a montré qu'il s'agissait là de gènes déjà connus. Par contre, les autres ADNc TSAP 4 à TSAP 8 ne montrent aucune homologie significative avec des gènes connus.

Pour l'ADNc TSIP 1 qui est inhibé dans son expression pendant l'apoptose, il ne montre aucune homologie avec des gènes connus.

20 Pour l'ADNc TSIP 2 qui est également inhibé dans son expression pendant l'apoptose, il montre une grande homologie avec le transcript S182 du gène AD3 impliqué dans les voies métaboliques de la maladie d'Alzheimer (Sherrington et al.) (figure 8).

25 Par conséquent, il est possible d'agir sur les voies métaboliques de la maladie d'Alzheimer en agissant sur les voies métaboliques p53 dépendantes.

La présente invention a donc également pour objet, à titre de médicament, un composé assurant l'expression cellulaire de TSIP 2 destiné au traitement de la maladie d'Alzheimer ainsi qu'à titre d'agent de diagnostic
30 pour la détermination de la prédisposition à la maladie d'Alzheimer, tout ou partie de la séquence de TSIP 2 à utiliser comme sonde nucléotidique ou comme amorce d'amplification ainsi qu'un antigène correspondant à tout ou partie des protéines codées par TSIP 2 ou les anticorps, notamment les anticorps monoclonaux correspondants, éventuellement après culture.

35 L'hypothèse que l'on peut faire sur ces gènes inhibés dans leur expression par le p53 sauvage est qu'ils peuvent coder pour des séquences oncogéniques qui seraient régulées en aval du processus de suppression

tumorale ou encore qu'il s'agit de protéines de structure ou du cytosquelette pour lesquelles la régulation en aval de l'expression est concomitante de la mort cellulaire par apoptose.

TSAP 1 est homologue à la phospholipase C bêta 4 de rat. La
5 séquence de TSAP 1 présente 100 % d'identité avec la PLC entre les
nucléotides 3967 et 3985 ; 82 % entre les nucléotides 3986 et 4116 et 85 % entre
les nucléotides 4070 et 4220 (figure 4). La PLC est connue pour être impliquée
dans la voie de signalisation des récepteurs de la tyrosine-kinase, et
pour catalyser l'hydrolyse du phosphatidylinositol-4,5-biphosphate en
10 diacylglycérol et inositol-1,4,5-triphosphate. Toutefois, la présente étude
suggère que la PLC est une cible en aval dans l'apoptose à médiation p53.

TSAP 2 montre des séquences conservées (92 % d'identité entre les
nucléotides 259 et 299 ; 100 % d'identité entre les nucléotides 418 et 458 et 92 %
d'identité entre les nucléotides 645 et 685) avec la protéine digitée au zinc
15 (ZFM 1) qui est localisée dans le locus Multiple Endocrine Neoplasia (MEN 1)
(figure 5). MEN 1 est un désordre dominant autosomal associé avec le
développement de tumeurs affectant le lobe antérieur des glandes pituitaires
et parathyroïdes et les cellules des ilots pancréatiques. Il est particulièrement
intéressant d'avoir mis en évidence qu'à la fois ZFM et une isoenzyme de PLC
20 sont colocalisés dans la même région chromosomique 11q13 contenant le
gène de susceptibilité à MEN 1. Chez la souris, les régions homologues sont
localisées sur le chromosome 19B. Le fait de trouver que TSAP 1 et TSAP 2 sont
activés en réponse à p53 peut suggérer que ces gènes appartiennent à une
voie de suppression des tumeurs plus globale et que p53 peut coopérer avec
25 MEN 1.

TSAP 3 est identique à Siah 1b. Ce gène est l'homologue chez les
vertébrés du gène *Drosophila* seven in absentia (*sina*). Le clone décrit
présente 94 % d'identité avec l'homologue murin (nucléotides 1496 à 1634)
(figure 6). Par analyse Northern blot en utilisant une sonde TSAP 3, on a pu
30 détecter une expression différentielle d'un messager de 1,9 kb de ce gène
(figure 2A). Ceci est confirmé en utilisant une seconde sonde correspondant à
la même région de la séquence *siah* 1b décrite (figure 2B). La figure 2C
montre la distribution tissulaire de ce gène en utilisant une sonde TSAP 3 qui
détecte à la fois l'ARNm de 1,9 et de 2,4 kb correspondant aux résultats
35 mentionnés précédemment lorsqu'une sonde *siah* est utilisée. L'hybridation
in situ montre que l'ARNm de TSAP 3 est induit rapidement 1 heure après
l'induction de l'apoptose (figure 3D). Son expression augmente après 2 et

4 heures (figures 3E et 3F). Dans les cellules qui sont entrées en mitose aucun signal n'est détecté.

Carthew et Rubin ont montré que *seven in absentia* est nécessaire pour le développement de l'oeil de la drosophile. D'autre part, des mutants de ce gène dans la drosophile montrent un rôle beaucoup plus général dans le développement. L'homologue murin est subdivisé en deux groupes *siah 1* et *siah 2* et ces protéines montrent un degré de conservation tout à fait inhabituel par rapport à *drosophila seven in absentia*.

Nos résultats ont montré que TSAP 3 / *siah 1b* est activé dans le programme de mort cellulaire dans les cellules M1 induites par le gène suppresseur de tumeur p53. Comme ce gène code pour une protéine digérée au zinc nucléaire, il pourrait être un facteur de transcription régulateur qui est en aval du signal de p53. Les résultats montrent également un lien direct entre les gènes concernant le développement chez la drosophile et une voie majeure de suppression tumorale.

EXEMPLE 2

En utilisant le fragment d'ADNc murin (TSAP 3), décrit ci-dessus, obtenu par analyse différentielle d'ARNm, on a constitué une sonde pour isoler un fragment de 1,1 kb d'une librairie d'ADNc humain qui ensuite a été expansé jusqu'à la région codante entière par une RACE-PCR.

La figure 7 montre l'ADNc et la séquence d'acides aminés du gène humain *sina* (TSAP 3).

Cette séquence code une protéine de 282 amino-acides avec un motif digité au zinc C3HC4. Cette protéine présente également des analogies avec des protéines capables de se fixer sur l'ARN. La séquence en amino-acides est très conservée entre la Drosophile, la souris et le gène humain (figure 7).

La distribution tissulaire indique que le *sina* humain est exprimé de façon ubiquitaire et code pour un ARNm de 2,3 kb et, dans le placenta, il existe un transcrit additionnel de 2,5 kb.

En analysant des YAC du CEPH et des librairies BAC par PCR, en utilisant des amorces *sina* humains spécifiques, on a pu isoler 8 YAC (350-1000 kb) et 2 BAC (100 et 125 kb).

La fluorescence par hybridation in situ (FISH) utilisant les clones YAC et BAC montre que le *seven in absentia* est localisé sur le chromosome 16q12-13, c'est-à-dire dans une région contenant les gènes suppresseurs de tumeurs candidat dans différents cancers, notamment : cancer du sein (9).

tumeur de Wilm's (10-12), syndrome de Laurence-Moon-Bard et-Biedl (13), syndrome de Beckwith-Wiederman (14).

Comme cela a été indiqué dans la demande de brevet français N° 95 15 146, on a trouvé que des transfectances stables de cellules M1 murines avec le mutant p53 sensible à la température montraient l'activation de seven in absentia après induction de l'apoptose à 32°C. Etant donné que le TSAP 3 murin a été isolé dans un modèle d'apoptose induit par le gène p53, il était logique d'approfondir l'analyse du gène TSAP3 (HUMSIAH) dans un modèle d'apoptose physiologique humain.

Ce modèle est décrit dans l'intestin où les cellules migrent du fond de la crypte vers la région apicale des vilosités où elles meurent par apoptose avant d'être larguées dans le lumen. Ces cellules en apoptose sont spécifiquement marquées par la technique TUNEL.

D'autre part, ces mêmes cellules sont positives par hybridation in situ pour le gène TSAP 3 (HUMSIAH) dans l'apoptose physiologique chez l'humain.

Enfin, afin d'investiguer l'implication du gène TSAP 3 humain dans la suppression des tumeurs, on a utilisé un modèle basé sur l'ensemble des gènes plutôt que sur un seul gène. Ce modèle repose sur les propriétés biologiques du parvovirus H-1.

Des recherches très complètes dans ce domaine ont montré sur les 20 dernières années que le parvovirus tue préférentiellement les cellules tumorales alors qu'il épargne leur contrepartie normale.

De façon à élaborer un modèle, on a fait l'hypothèse suivante : s'il était possible de sélectionner, à partir d'une tumeur qui soit sensible à l'effet cytopathique du parvovirus H-1, les cellules qui étaient résistantes, cette résistance pourrait être due à un changement de leur phénotype malin. Ceci a pu être démontré pour les cellules KS sélectionnées à partir des cellules érythro-leucémiques K562 humaines. Tandis que les cellules parentales K562 sont sensibles à l'effet cytopathique du parvovirus H-1, les cellules KS, elles, sont résistantes. Ces cellules résistantes réexpriment le type sauvage de p53 et ont un phénotype supprimé à la fois in vitro et in vivo.

Pour confirmer ces observations sur d'autres cellules, on a sélectionné, à partir d'un monoclonal d'une leucémie monocyttaire U937 humaine, les cellules filles US3 et US4. Ces clones sont résistants à l'effet cytopathique des parvovirus H-1 et montrent une réversion du phénotype malin in vivo. L'analyse de marqueurs de surface pour 20 cellules, indique

qu'il n'y a pas de déplacement dans le stade de différenciation entre U937 et les clones US indiquant que la suppression du phénotype malin n'est pas due à une différenciation terminale.

Ni les cellules K562 ni les cellules U937 n'expriment p53. Par
5 contraste aux cellules KS qui réexpriment p53, les cellules US3 et US4 ne réexpriment p53. Toutefois, on a pu mettre en évidence le fait que les cellules US3 et US4 montraient l'activation de WAF-1 par rapport aux cellules parentales malignes U937. Une telle activation de WAF-1 dans une voie indépendante de p53 alternative a été récemment décrite et les résultats
10 actuels montrent que les clones US3 et US4 utilisent, semble-t-il, cette voie alternative WAF-1.

Le gène sina est activé par le type sauvage p53 inductible dans les cellules M1 de même que dans les cellules KS qui réexpriment le type sauvage p53.

15 Tandis que les cellules parentales U937 expriment très légèrement l'ARNm de sina, il est activé dans les clones filles US3 et US4 qui ont une réversion de leur phénotype malin et qui réexpriment p21^{waf-1}.

De façon intéressante, sina est activé dans les cellules qui deviennent apoptotiques, comme cela est montré par un double marquage
20 utilisant une sonde sina pour hybridation in situ combinée avec un essai TUNEL.

Ceci permet de démontrer que le gène sina humain qui est très conservé dans la phylogénie joue un rôle dans l'apoptose et la suppression tumorale.

25 De façon encore plus importante, sina se situe au croisement des voies de p53 et de WAF-1.

En outre, en utilisant le modèle de U937 et US3 et US4, on a pu montrer un lien fonctionnel pour les molécules suppresseurs en utilisant un modèle biologique global qui permet la comparaison à des niveaux
30 moléculaires entre les cellules malignes parentales et les cellules filles directement dérivées. Ces expériences indiquent qu'il n'est pas nécessaire de transférer les gènes suppresseurs de tumeur humains spécifiques de façon à leur conférer le phénotype suppresseur, mais que la réversion tumorale est sous le contrôle d'un système de régulation qui est toujours présent dans le matériel génétique des cellules tumorales bien qu'il soit nécessaire de le
35 réactiver.

TABLEAUCARACTERISTIQUES DES CLONES

5		<u>Clone à expression différentielle</u>	<u>Amorces 3' et 5' *</u>	<u>Taille de l'ARNm en kb</u>	<u>Homologie</u>
		TSAP 1	T11GC-16	2,0 et 4,5	PLC #
		TSAP 2	T11GC-5	5,9	MEN1 §
10		TSAP 3 (IDS N° 3)	T11CG-4	1,9	siah 1b ¶
		TSAP 4	T11GC-6	5,0	Non
		TSAP 5	T11CG-5	1,2	Non
		TSAP 6	T11AG-1	2,8	Non
		TSAP 7	T11GC-16	> 8,0	Non
15		TSAP 8	T11GC-6	> 10,0	Non
		TSIP 1	T11CG-8	3,0	Non
		TSIP 2	T11AA-5	3,1	AD3 *

20 * Les chiffres et les séquences des amorces en 5' correspondent à ceux rapportés par Bauer et al. (4)

Rat phospholipase C-bêta 4 ARNm (RATPHOSCB)

§ ARNm humains (HUMMEN1C; HUMZEN1C; HUMZEN1A; HUMMEN1A)

¶ siah-1B ARNm (MINSIAH1B)

25 * AD3, transcript S182 ARNm murin (homologue S182 ARNm humain) (Sherrington et al.).

- (A) LONGUEUR:

- 5 (C) NOMBRE DE BRINS: simple

- (D) CONFIGURATION : linéaire

- (ii) TYPE DE MOLECULE : ADNc

- (ix) CARACTERISTIQUE:

- (A) NOM/CLE: TSAP 1

- 10 (B) EMPLACEMENT:

- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 1 :

10.

TGATCACGTAC

rat?LC

CTTCCTCTACTTAACAATTTGACTATTGAATTTCTTTGGCCCAACCAAAAGTAGCTATGTAC

1970	1980	1990	2000	2010	2020
------	------	------	------	------	------

TSAP:

[illegible][illegible]

706

ACACACACACACACACACACACACACA-----CACACACACACACACAGAATT

4030 4040 4050 4060

80 90 100 110 120 130

TSAP:

CCCCTATTCCTGACAGGCAGAGTTGAATCATGATATATGGCTTAAACATGTTTGCTATGA

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 104

rat?LC

CCCCCTATTCTGACAGGCAGAGTTGAACCATATTCACAACTTAAACATGTTGCCTAGGC

4070	4080	4090	4100	4110	4120
------	------	------	------	------	------

35

```

                140      150      160      170      180      190
TSAP 1  GACAGCATCACAAGCCAGTGGGCTTGGTGATAACAACCTCTGCTTTGTGGTGCATTAGGAC
          .....
5  ratPLC GACAGCATCACAAGCCAGTGGGCTTGGTGATAACAACCTCTGCTTTGTGGTGCATTAGGAC
          4130      4140      4150      4160      4170      4180

```

```

                200      210      220      230
10 TSAP1 1 ATTTTGTAGCTGCTGCTGCTGCAAA-AAAAATAAGAGCCG
          .....
ratPLC ATGTTTCGAGCTGCTGCTG--GAAAAGGAAAATTAGTGCATTAGTACTTAAATGCCAAGCC
          4190      4200      4210      4220      4230      4240

```

15 INFORMATION POUR LA SEQ ID N° : 2

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR:

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

20 (D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

(A) NOM/CLE: TSAP 2

(B) EMPLACEMENT:

25 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 2 :

TSAP2

```

10      20      30      40      50      60
30 TSAP2 CCTTGGAAACCAATCTACAACAGCGAGGGGAAGCGGCTTAACACCTCGAGAGTTCCGTACCC
          .....
humfmlc.seq CCCCTGAGCCCATCTACAATAGCGAGGGGAAGCGGCTTAACACCTCGAGAGTTCCGTACCC
          250      260      270      280      290      300

```

```

              70      80      90      100      110      120
TSAP2          GCAAAAAAAAAAATCTCTTGTGTTTTCTAAGCTTTTCCTGTGCTAGGGAAAGATCACT
              : : : : :
5  hum2fmlc.seq GCAAAAAGCTGGAAGAGGAGCGGCACAACCTCATCACAGAGATGCTTGCACTCAATCCGG
              310      320      330      340      350      360

              130      140
10 TSAP2          AACTCCCTGGTTATAGATTGCTT

              370      380      390      400      410
              hum2fmlc.seq ATTCAAGCCACCTGCAGATTACAAACCTCCAGCAACAGTGTGAGTGAT

15  (i) INFORMATION POUR LA SEQ ID N° : 3
      (i) CARACTERISTIQUES DE LA SEQUENCE:
            (A) LONGUEUR:
            (B) TYPE: nucléotide
            (C) NOMBRE DE BRINS: simple
20      (D) CONFIGURATION : linéaire
      (ii) TYPE DE MOLECULE : ADNc
      (ix) CARACTERISTIQUE:
            (A) NOM/CLE: TSAP 3
            (B) EMPLACEMENT:
25  (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 3 :

      TSAP3
      10
      TSAP3 3
30
      mmshah1b.seq TTGTAAAATATTTCTGAACTTTGTATTTGTTGTAGATTGATTGTATTGTTGACAATTTTT
              1450      1460      1470      1480      1490      1500

```



```

      20      30      40      50      60      70
TSAP 3      CCGGGTGGGGGTGTGCCTGCACACATGCGTGCACGTGTGTGCTTGGTTTTCTTTAACA
      .....
5  mmsiah1b.seq CCGGGTGGGGGTGTGCCTGCACACATGCGTGCACGTGTGTGCTTGGTTTTCTTTAACA
      1510      1520      1530      1540      1550      1560

      80      90      100      110      120      130
10 TSAP 3      GCCATCTACGTGTCATAGCCCACTGTTTTCCCCTTGTGAGTCAACACATAGTCTGCTGT
      .....
      mmsiah1b.seq GCCATCTACGTGTCATAGCCCACTGTTTTCCCCTTGTGAGTCAACACATAGTCTGCTGT
      1570      1580      1590      1600      1610      1620

15
      140
      TSAP3      GGTTCGGGTTTGCT
      .....
20  mmsiah1b.seq GGTTCGGGTTTGCTTTTGGTTTTGGATGCTGTGTGATTTGATAATTTTATTCTA
      1630      1640      1650      1660      1670      1680

```

25

30

35

INFORMATIONS POUR LA SEQ ID N° : 4

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR:

(B) TYPE: nucléotide

5 (C) NOMBRE DE BRINS: simple

(D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

(A) NOM/CLE: TSAP 4

10 (B) EMPLACEMENT:

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 4 :

TSAP4

AACTCCGTCG TGGGTGTGGG GACCTAATTC CTTATATTTT TACAACAAGC ACTGTACAAA 50
 15 CTGTGCCTTT CCCTAATGCA GTTATACTAT TTCCATTAAG ATGGGTAACC TTAGTTAAGG 120
 CTTTATATTC ACTGCCATGG GTAGGAATGC TCACGGTGAA TGGGCGAACT TGTGATGGAA 180
 GAGGCCCTCA TTTTCAGTTG GC 202

20 INFORMATIONS POUR LA SEQ ID N° : 5

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR:

(B) TYPE: nucléotide

(C) NOMBRE DE BRINS: simple

25 (D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

(A) NOM/CLE: TSAP 5

(B) EMPLACEMENT:

30 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 5 :

TSAP5

	TAACAAGGAT ATTCAGGTTT GGGATTGCTT TCCTAAGCGA TGATCTCAAC CTCCACGTGG	60
5	AACTGATTTC CCAAGGGACA GAAATGGTCT TTGATCTTTC TGAACCACTT GTCTTCAAAC	120
	TCTTTGGAGG ACGCAACCAC CATGGCAGTC AGGGCTCCGG GGGCCACACA CTTACCTCC	180
	GAATGAAGCT CCTCTTTTAT CTTTCTGGG ACAATGTCTT CCCCCATAGC CTCCTCCATC	240
	AACAGCAAAG TACCTTCCCT AAAGTTGAAG TCCTTCACTT TCCCTGCAAT TTCTGTCTGA	300
10	GTCTCAAGT TCTTCTCCAA CGCGAATGAT GTTTGCTGAG ACTGGGCGAG CTGAAGCAGG	360
	AGCCTGGCGC GGAGCAAAAA GCGCGATGCT TTCCTCCGAG CCTCCATCTG TGCTCTTCC	420
	CTCCGCTTGG CCAGGGAAGG CATATTCTC CTGAGCACTA CCACTCGCTT CCACGGAGAG	480
	CAGTGCATTC TCAGGCAAGG TCGTGGGCAA AGACAAAAGA GAGCCTGTTC CCGAGTGTAC	540
15	AGAGGAGGGA CCGACGGCCT TGTCATTGA GGCAGAACTC TTCTGTCCCT GCGGTGACAC	600
	CCTGCTGGCA GCGCGGGCCC TGGACTCAGG TATGCTCTG CCAGCTTACA CCAGCTCCAC	660
	GGGTTGAGCG GGTGCAAAGC AATCAGCTTG TGCAGGCAGA AGATCGTGTG CTCCCGGCTC	720
	TGCAGGCTGG AAAAGACGGC CAGGTGGAGG TGGAGCAGCA CGGTCAGATG GTCTGTCTTG	780
20	GTGGCTTTGC TTTCGAAGTC TGCGGCCATC TCCAGCGCCT CCGCATGCCT CCGAAGTGAG	840
	CCAGACACCG AGCCTGGCCT TCTTGGACAT CCCTTTTCAT GGCAAAATTA GTAGATGCTA	900
	ATGTTCCGAG ATATGSAGTA TTCCTGCAGG GCTTTCTCTT ATTCTGTCTG TCTGTAGCCC	960
25	AGGTCCCTTC TGAATTTCTT GAGAGTGAGA ACTTCAATAT CGTCACTACA TTCTGTCTCT	1020
	TCATAAAACC ATGCGGCTCG CAGAGCTTGG CGCGGTAGGG GGAGGGCGGC TCGGGCGGGC	1080
	GCTCCGGCCT CTGCTCGAAC ACCGAGTCCT CAAATTCGCC GCGCAGCACC CAGCATCCGG	1140
	TCTCCATCGC GCGGAAGTGC AACTGGACCT CGAAACGAGG CGACACCTAG AGCGACGCCC	1200
30	ATCAGCCAGC CTCCAAAGCG CCGACAGCA GCGCGCCCAA GGCTGCCGAG GCAAGGTAGA	1260
	GACCTGCCCC GCGGGCCGCT CGAGCCCTAT AGTGAGTCTT ATTAGGATGG	1310

INFORMATIONS POUR LA SEQ ID N° : 6

(i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR:

(B) TYPE: nucléotide

5 (C) NOMBRE DE BRINS: simple

(D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

(A) NOM/CLE: TSAP 6

10 (B) EMBLACEMENT:

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 6 :

TSAP6

15	GTGAGTACAT ATCACATGTA TGSGGTGTCA TTCTGAGTAT GTCAGTTTAC ACCTGCATCC	60
	CAGGAATTAG GATCTCAGCC ACCCAGCAT ATATCATCAC CTCCTGTTC AGCATCCAGA	120
	AAAGAGACCC GAACCCAGCT CAGGCCCCC ACAAGCCATC TCCACTTCCA GGGCTCACA	180
	CGTGGCTTGT TTTCTCCCC TGTGTGTGGT GCGCGGACAG CATGAAGTTG ACAGCCCCAT	240
20	CTTTCTCCCA GGGCTGCGG ATCTTGCTGA GTCTGCGGT TGAGGCAGG CAGGAGGAAG	300
	AGGCTCTTGG CCAGGATGAT TCACACAGGG GCAGGGAGCA GCTGAGTGT GGAATGTGGG	360
	GCGGGCAGGT AGAAGTTGAT AGTGTTTTT CTTCAAAAG GCAGGGGTCC AGGCTAGGT	420
	GAGTGTGTGC ATTGTGCTGA GTATCAGGGC CACGAAGCCC AGTGTGGACT GCACGAAGCT	480
25	GAATCTCTTC CAGTTGAGGG AATTAGCAAT GCAGGGGAGC GAGGTGACAG CCAGCAGCGA	540
	CAACATGCCC AGGGCCAGCA CAGCCAGGGA CAGGTATATC TCCATCCTCC AGACTTCTTC	600
	CTCAGCCCAG AGGCGGCTCT TGTGCGCCAG GACCTGCTTC ACAGCCAGAT TGACCAGGTC	660
	GTAGGCGGTG GGAGCGGCGC AGCGGCAGGC AGAAGCTGTA GAGAGCGTCC AGCATCCGA	720
30	AGAAGAAGCT GAGCAGCCCC ATCTGCTTGC GATGCTGCAG CCAAGGTGTC AGCAGTCTC	780
	GGAAGCGCTG GTACTTGCTC CCCCCTCCGA GCTGAAGGCG ACCTGCCAGC ACACCGGCA	840
	GGTACACTAG GGACAGCAGC ACATAAGCCA CACAGGCTAG TGTGTTGTTG ACCACAGACA	900
	AGGGCATCTT GTAAAGCTG TTCTCATCTT TCCGAATGTA TGGTGTANA ACCTCCCGGA	960
35	TGAAATTGTA GGTGTANAAN CACACAAAGA CCCCAGTCCC CAGGAAGGTC GGGGCTTCC	1020

AGAATGGAAG GAAGCNCAGG GGTTCNGCTT CTACCTCCCT CAGTGAAGGC CAGGATCCA 1080
 TNTCCAGGGG TTNAACCAT NGGGCGTGCA TCTCTGAAA TGCTCCTTG GNTTCTGCTH 1140
 GATCANTGCA AATAACNCCT GCCTGTTCEN TCCCTTGGGG CCACCCCTNTT GGGGCCATGC 1200
 5 CAA 1203

INFORMATIONS POUR LA SEQ ID N° : 7

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR:
- 10 (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- (D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

- 15 (A) NOM/CLE: TSAP 7
- (B) EMPLACEMENT:

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 7 :

TSAP7

20 GGGCATCCAG TCATTCTTTA TTTCAGTGTG TGAAGGCTG CTACGCATTT TCCGCCAAT 50
 TAATTTTAA TCCATTTTCA AACGAGCCTT TACTGTGGCC TTTTCTGCTA TTTTGCATAT 120
 ATGTTAGCAC GTGTGCATAG 140

25 INFORMATIONS POUR LA SEQ ID N° : 8

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR:
- (B) TYPE: nucléotide
- (C) NOMBRE DE BRINS: simple
- 30 (D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: TSAP 8
- (B) EMPLACEMENT:
- 35 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 8 :

TSAP8

CACGTNAAAG TACCACATCC NCCCCCATTS GTAGATATTG ANAGAGTATA TANATAGGNC 60
 GAAGCACAAT CTCTTCCCTT CCTNTGTACA CCTCANACCC AGTGACTTCC NACCNAAAGCN 120
 5 CNTGANTGTN TTTGTNGATA TGAGTGTCTG NGTGTGTGNA TNTCCGTCTC ACATGTATGG 180
 GACGACCNAC CCCACCCCCA GCGGCCTTCA NGCACAATNG AGGACCCCTA TNGTGGATAC 240
 GNGCATCGGT_AAAANAGC 257

INFORMATIONS POUR LA SEQ ID N° : 9

- 10 (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR:
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - (D) CONFIGURATION : linéaire
- 15 (ii) TYPE DE MOLECULE : ADNc
- (ix) CARACTERISTIQUE:
- (A) NOM/CLE: TSIP 1
 - (B) EMBLACEMENT:
- (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 9 :

20 TSIP1

GGAGGGGGTG TAGCTTTCTG TTTAGTTATC ACTCTGAGGT GGTGAGGTCA CAGAGAAAGG 60
 ACTTAATTGG GAGGTCATC TGATTCGGGG CATCTTCTCT CCTTTACCA A 111

25 INFORMATIONS POUR LA SEQ ID N° : 10

- (i) CARACTERISTIQUES DE LA SEQUENCE:
- (A) LONGUEUR:
 - (B) TYPE: nucléotide
 - (C) NOMBRE DE BRINS: simple
 - 30 (D) CONFIGURATION : linéaire
- (ii) TYPE DE MOLECULE : ADNc
- (ix) CARACTERISTIQUE:
- (A) NOM/CLE: TSIP 2
 - (B) EMBLACEMENT:
- 35 (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 10 :

TSIP2

	CACCGGTGAGACCTCTAGGGCGGGGCTAGGACGACCTGCTCCGTGGGCGCGAGTATTC	60
	GTCGGAAACAAAACAGCGGCAGCTGAGGCGGAAACCTAGGCTGCGAGCCGGCCGCCCGGG	120
5	CGCGGAGAGAGAAGGAACCAACACAAGACAGCAGCCCTTCGAGGTCTTTAGGCAGCTTGG	180
	AGGAGAACACATGAGAGAAAGAATCCCAAGAGTTTTGTGTTCTTTGAGAAGGTATTTCT	240
	GTCCAGCTGCTCCAATGACAGAGATACCTGCACCTTTGTCTACTTCCAGAATGCCAGAG	300
	TGTCTGAGGACAGCCACTCCAGCAGCGCCATCCGGAGCCAGAATGACAGCCAAGAACGGC	360
10	AGCAGCAGCATGACAGGCAGAGACTTGACAACCCCTGAGCCAATATCTAATGGGCGGCCCC	420
	AGAGTAACTCAAGACAGGTGGTGGAACAAGATGAGGAGGAAGACGAAGAGCTGACATTGA	480
	AATATGGAGCCAAGCATGTCATCATGCTCTTTGTCCCCGTGACCCCTCTGCATGCTCGTCG	540
	TCGTGGCCACCATCAAATCAGTCAGTTCTATACCCGGAAGGACGGTCAGCTAATCTACA	600
	CCCCATTCACAGAAGACACTGAGACTGTAGGCCAAAGAGCCCTGCACTCGATCCTGAATG	660
15	CGGCCATCATGATCAGTGTCATTGTCAATTATGACCATCCTCCTGGTGGTCTCTGTATAAAT	720
	ACAGGTGCTACAAAGTCATCCACGCCCTGGCTTATTATTTTCATCTCTGTTGTTGCTGTTCT	780
	TTTTTTCTGTTTCAATTTACTTAGGGGAAGTATTTAAGACCTACAATGTGCGCCGTGGACTACG	840
	TTACAGTAGCACTCCTAATCTGGAAATTTTGGTGTGGTGGGATGATTGCCATCCACTGGA	900
20	AAGGCCCCCTTCGACTGCAGCAGGCGTATCTCATTATGATCAGTGCCCTCATGGCCCTGG	960
	TATTTATCAAGTACCTCCCCGAATGGACCGCATGGCTCATCTTGGCTGTGATTTCACTAT	1020
	ATGATTTGGTGGCTGTTTTATGTCCCAAAAGGCCCACTTCGTATGCTGGTTGAACAGCTC	1080
	AGGAAAGAAATGAGACTCTCTTTCCAGCTCTTATCTATTCTCAACAATGGTGTGGTTGG	1140
	TGAATATGGCTGAAGGAGACCCAGAAACCCCAAGGAGGTACCCAAAGAACCCCAAGTATA	1200
25	ACACACAAAAGAGCGGAGAGAGAGACACAGGACAGTGGTTCTGGGAACCATGATGGTGGCT	1260
	TCASTGAGGAGTGGGAGGCCCAAGAGACAGTCACCTGGGGCCTCATCGCTCCACTCCCCG	1320
	AGTCAAGAGCTGCTGTCCAGGAACCTTTCTGGGAGCATTTCTAACGAGTGAAGACCCGGAGG	1380
	AAAGAGGAGTAACTTTGGACTGGGAGATTTTCAATTTCTACAGTGTCTGGTTGSTAAGG	1440
30	CCTCAGCAACCGCCAGTGGAGACTGGAACACAACCATAGCCTGCTTTGTAGCCATACTGA	1500
	TCGGCCTGTGCCCTTACATTACTCCTGCTCGCCATTTTCAAGAAAGCGTTGCCAGCCCTCC	1560
	CCATCTCCATCACCTTCGGGCTCGTGTCTACTTCGCCACGGATTACCTTGTGCAGCCCT	1620
	TCATGGACCAACTTGCATTCCATCAGTTTTATATCTAGCCCTTTCTGCAGTTAGAACATGG	1680
	ATGTTCTCTCTTTTATTATCAAAAACACAAAAACAGAGAGCAAGCCCGAGGAGGAGACTG	1740
35	GTGACTTTCCTGTGTCTCTCAGCTAAACAAAGGCAGGACTCCAGCTGGACTTCTGCAGCTTC	1800
	CTTCCGAGTCTCCCTAGCCACCCGCACTACTGGACTGTGGAAAGGAAGCGTCTACAGAGGA	1860

INFORMATIONS POUR LA SEQ ID N° : 11

(i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR:
 (B) TYPE: nucléotide
 5 (C) NOMBRE DE BRINS: simple
 (D) CONFIGURATION : linéaire

(ii) TYPE DE MOLECULE : ADNc

(ix) CARACTERISTIQUE:

- (A) NOM/CLE: TSAP 3 humain
 10 (B) EMLACEMENT:

(xi) DESCRIPTION DE LA SEQUENCE: SEQ ID N° 11 :

TSAP3 humain

```

m s r q t a t a l p t g t s k c p p s q
atgagccgctcagactgctacagcattacctaccggtacctcgaagtgtccattatccag
10      20      30      40      50      60

r v p a l t g t t a s n n d l a s i f e
agggcgccctgcccctgactggcacaactggcattcccaacaaatgacttggcgagctcttttgag
15      70      80      90      100     110     120

c p v c f d y v l p p i l q c q s g h l
tgctcagctctgttttgactatgtgttaccgcccattctttcaatgtcagagtggtccatttt
130     140     150     160     170     180

v c s n c r p k l t c c p t c r g p l g
gtttgttagcaactgttcgccccaaagctccacatgtgtgtccaaacttggccggggccctttggga
20     190     200     210     220     230     240

s i r n l a m e k v a n s v l f p c k y
tcatttcgcaacttgggtatggagaaagtgggttaattcagtaattttttcccttgtaaatat
250     260     270     280     290     300

a s s g c e i t l p h t e k a d h e e l
gggtttttctggatgtgaaataactctgtccacacacagaaaaagcagaccatgaagagcttc
25     310     320     330     340     350     360

c e f r p y s c p c p g a s c k w q g s
tgtgagttttagggcctttattctctgtccgtgcccctgggtgcttctctgtaaatggcaaggcttc
370     380     390     400     410     420

l d a v m p h i m h q h k s i t t l q g
ctggatgctgttaattgccccatctgatgcacagcataagtcattacaacccctacagggga
430     440     450     460     470     480

e d i v f l a t d i n l p g a v d w v m
gaggatatagttttttctgtacagacatttaattcttctgtgtgtgtgactgggtgatg
30     490     500     510     520     530     540

m q s c f g f h f m l v l e k q e k y d
atgcagtcctgtttttgggttttcaattcatgttagctcttagagaaacaggaaaaaacgat
550     560     570     580     590     600

g h c q c f f a i v q l i g t r k q a e n
gggtaccagcagttttttcgcaatcgtacagctgataggaaacacgcaagcaagctgaaat
35     610     620     630     640     650     660

```


r s i h e g i a t a i m n s d c i v i d
cgatctatttcattgaaggaattgcaacagccattatgaatagcgactgtctagctcttgac
730 740 750 760 770 780

atgtgttgaaatggcaatcaaacatctctctggccagtggtttaaacttcagtttcacaga
850 860 870 880 890 900

gaagggccaatzaaaagaaagactgctaaatcacaggaaacagttcccatgtagttaacacttaa
970 930 990 1000 1010 1020

tatatttataaaataagtcgaacagctaaaccactgaaaaaatatatgtatatatccacccaaga
1030 1040 1050 1060 1070 1080

ctgtatgattgattctgtattgtctgaaaaagctctgtctcttcgcgtgggaagctgtctgtctctctgtctgtctg

1210 1220 1230 1240 1250

ctggcacaattcttataaaccttcagttcttcacccaataaattcgacctcttcgttaattctag
1330 1340 1350 1360 1370 1380

1390 1400 1410 1420 1430 1440

aaaaatttttaaaacagcctatttttggtgaaccatgagttcccagaaagttaaaggtgac
 1510 1520 1530 1540 1550 1560

a c c c g g a a a a a t a a t c c a a a a g c c c t a t t t a a a g c c a c c t a t a a g g t g c c c c c c t t t c c t g
 1570 1580 1590 1600 1610 1620

gattttataaaatacttgcaaatccaggctcttctgtttccctttttccagatatccctggaca
1590 1700 1710 1720 1730 1740

aaacacatctcttcaaaattctgtctctgtatctattgggtctcttcagagaagaaggcatctgtca
1750 1760 1770 1780 1790 1800

gccccggccccctataatccca
1370 1880

REFERENCES BIBLIOGRAPHIQUES

- (1) Liang P. & Pardee A.B. (1992) Science, 257, 967-971.
- (2) Don R.H., Cox P.T., Wainwright B.J., Baker K. & Mattick J.S. (1991) Nucl.
5 Acids Res., 19, 4008.
- (3) Yonish-Rouach E., Resnitzky D., Lotem J., Sachs L., Kimchi A. & Oren M.
(1991) Nature 352, 345-347.
- (4) Bauer D., Muller H., Reich J., Riedel H., Ahrenkiel V., Warthoe P. &
Strauss M. (1993) Nucl. Acids Res. 21, 4272-4280.
- 10 (5) Sambrook J., Fritsch E.F. & Maniatis T. (1989) Molecular Cloning: a
laboratory manual.
- (6) Okamoto K. & Beach D. (1994) EMBO J., 13, 4816-4822.
- (7) Angerer L. & Angerer R.C. (1991) Methods in cell biology : functional
organization of the nucleus, 35, 37-71.
- 15 (8) Linares-Cruz G., Rigaut J.P., Vassy J., De Oliveira T.C., De Cremoux P.,
Olofsson B. & Calvo F. (1994) J. Microsc., 173, 27-38.
- (9) Bieche I. and Lidereau R., Genes Chromosomes and Cancer 14, 227-251
(1995).
- (10) Wang-Wuu S., Soukup S., Bove K., Gotwals B. and Lampkin B., Cancer
20 Research 50, 2786-2793 (1990).
- (11) Maw M.A. et al., Cancer Research 52, 3094-3098 (1992).
- (12) Austruy E. et al., Genes, Chromosomes and Cancer, 14, 285-294 (1995).
- (13) Kuytek-Black A.E. et al., Nat. Genet 5(4) 392-396 (1993).
- (14) Newsham I. et al., Genes Chromosomes and Cancer 12(1), 1-7, (1995).
- 25 (15) Sherrington et al., Nature, vol. 375, p. 754-760 (1995).

REVENDICATIONS

- 1) Séquence nucléotidique correspondant à un gène comportant :
 - (a) une séquence selon l'une des IND.SEQ 4 à 11 ou
 - 5 un gène équivalent qui comporte :
 - (b) une séquence s'hybridant avec l'une des séquences selon (a),
 - (c) une séquence présentant au moins 80 % d'homologie avec (a) ou (b), ou
 - (d) 10 une séquence codant pour une protéine codée par un gène selon (a), (b) ou (c) ou pour une protéine équivalente.
- 2) Séquence selon la revendication 1, caractérisée en ce que l'expression cellulaire du gène est induite lors de l'apoptose cellulaire.
- 3) Séquence nucléotidique correspondant à un gène comportant :
 - (a) 15 une séquence selon l'une des IND.SEQ 1 et 3 ou un gène équivalent qui comporte :
 - (b) une séquence s'hybridant avec l'une des séquences selon (a),
 - (c) une séquence présentant au moins 80 % d'homologie avec (a) ou (b), ou
 - (d) 20 une séquence codant pour une protéine codée par un gène selon (a), (b) ou (c) ou pour une protéine équivalente,
 caractérisée en ce que l'expression cellulaire du gène est induite par la suppression tumorale.
- 4) Séquence nucléotidique correspondant à un gène comportant :
 - (a) 25 une séquence selon l'une des IND.SEQ 2 ou un gène équivalent qui comporte :
 - (b) une séquence s'hybridant avec l'une des séquences selon (a),
 - (c) une séquence présentant au moins 80 % d'homologie avec (a) ou (b), ou
 - (d) 30 une séquence codant pour une protéine codée par un gène selon (a), (b) ou (c) ou pour une protéine équivalente,
 caractérisée en ce que l'expression cellulaire du gène est induite par l'apoptose cellulaire.
- 5) Séquence selon l'une des revendications 1 à 4, caractérisée en ce que l'expression cellulaire du gène est induite par p53.
- 35 6) Séquence selon la revendication 2 ou 4, caractérisée en ce que l'apoptose cellulaire est induite par p53.

7) Séquence selon l'une des revendications 1 à 6, caractérisée en ce qu'elle est choisie parmi TSAP 1 à TSAP 8 et TSAP 3 humain ou un gène équivalent.

5 8) Séquence selon la revendication 1, caractérisée en ce que l'expression cellulaire du gène est inhibée lors de l'apoptose cellulaire.

9) Séquence selon la revendication 8, caractérisée en ce que l'apoptose cellulaire est induite par p53.

10 10) Séquence selon l'une des revendications 1 et 8 et 9, caractérisée en ce qu'elle est choisie parmi TSIP 1 et TSIP 2 ou un gène équivalent.

11) Vecteur d'expression cellulaire d'une séquence selon l'une des revendications 1 à 10.

12) Vecteur d'expression selon la revendication 11, caractérisé en ce qu'il s'agit d'un vecteur viral.

15 13) Vecteur selon la revendication 12, caractérisé en ce qu'il s'agit d'un adénovirus, d'un rétrovirus, d'un virus herpès ou d'un poxvirus.

14) Vecteur selon la revendication 11, caractérisé en ce qu'il s'agit d'un vecteur à acide nucléique nu.

20 15) Vecteur selon l'une des revendications 11 à 13, caractérisé en ce qu'il comporte une séquence assurant le ciblage et/ou l'expression spécifique des tissus ou organes.

16) Cellule transformée par un vecteur d'expression selon l'une des revendications 11 à 15.

25 17) Protéine pouvant être obtenue par culture de cellule transformée selon la revendication 16 et codée par la séquence selon l'une des revendications 1 à 10.

18) A titre de médicament, un vecteur selon l'une des revendications 11 à 15 ou une protéine selon la revendication 17.

30 19) A titre de médicament, un composé assurant l'expression cellulaire d'au moins une des séquences nucléotidiques selon l'une des revendications 1 à 7 ou de leurs produits.

20) A titre de médicament selon la revendication 19, un vecteur nucléotidique assurant l'expression cellulaire de ladite séquence.

35 21) A titre de médicament, un composé assurant l'inhibition de l'expression cellulaire d'au moins un gène cellulaire selon l'une des revendications 1, 8 à 10 ou de leurs produits.

22) A titre de médicament selon la revendication 21, un nucléotide activé assurant le blocage de la séquence nucléotidique.

23) A titre de médicament selon la revendication 21, un anticorps monoclonal dressé contre la ou les protéines codées par la séquence
5 nucléotidique.

24) A titre de médicament destiné au traitement du cancer, un médicament selon l'une des revendications 18 à 23.

25) A titre de médicament destiné au traitement de la maladie d'Alzheimer, un médicament selon l'une des revendications 18 à 23.

10 26) A titre d'agent de diagnostic notamment pour la détermination de la prédisposition et le suivi des cancers, tout ou partie des séquences selon l'une des revendications 1 à 10 à utiliser comme sonde nucléotidique ou comme amorce d'amplification.

15 27) A titre d'agent de diagnostic notamment pour la détermination de la prédisposition et le suivi des cancers un antigène correspondant à tout ou partie des protéines codées par la séquence selon l'une des revendications 1 à 10 ou les anticorps correspondants.

20 28) A titre d'agent de diagnostic notamment pour la détermination de la prédisposition et le suivi de la maladie d'Alzheimer, tout ou partie des séquences selon l'une des revendications 1, 5, 7 à 10 à utiliser comme sonde nucléotidique ou comme amorce d'amplification.

25 29) A titre d'agent de diagnostic notamment pour la détermination de la prédisposition et le suivi de la maladie d'Alzheimer un antigène correspondant à tout ou partie des protéines codées par la séquence selon l'une des revendications 1, 5, 7 à 10 ou les anticorps correspondants.

30) A titre d'agent antiviral, un médicament selon la revendication 20.

31) Modèle pour la mise en évidence de médicament anti-cancéreux, des cellules selon la revendication 16.

30 32) A titre de perfectionnement de la méthode de Liang et Pardoll le fait d'utiliser une diminution en palier lors de l'amplification PCR.

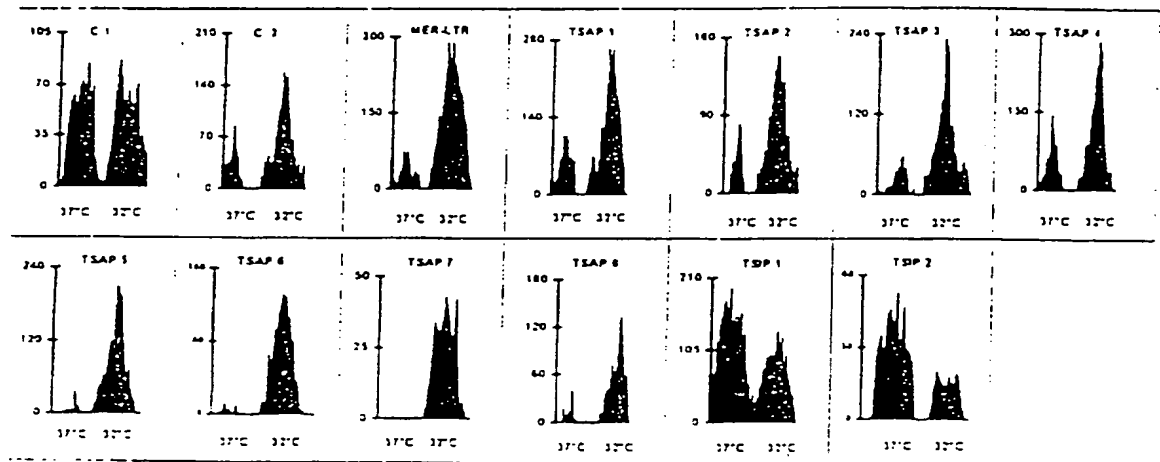


FIG. 1

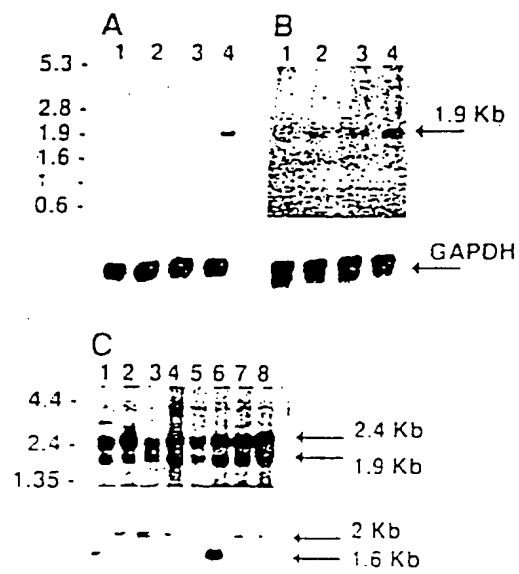


FIG. 2

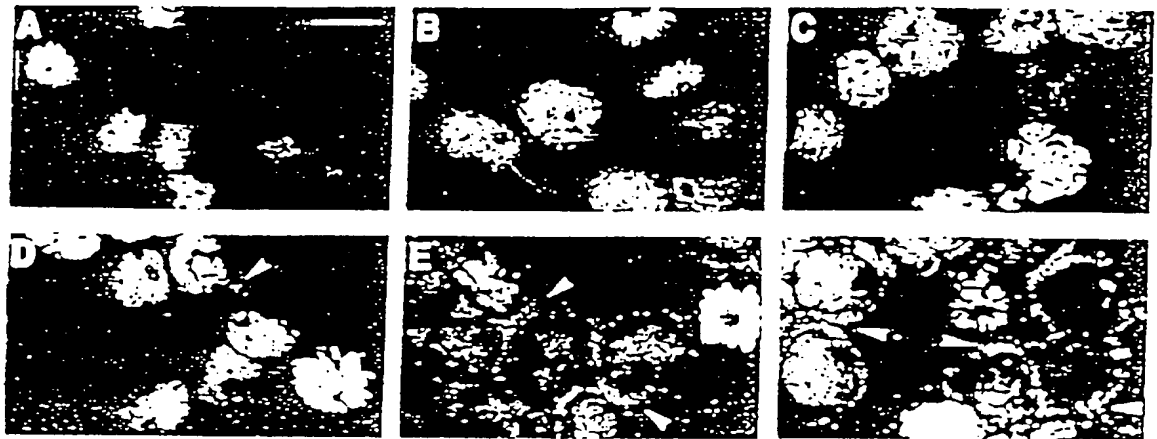


FIG. 3

TSAP 2

```

10      20      30      40      50      60
TSAP2    GCTTGGAAACCAATCTACAACAGCGGGAAGCGGCTTAACACTCGAGAGTTCCGTACCC
          :: :: ::::: ::::: ::::: ::::: ::::: ::::: :::::
humzfm1c.seq CCCCTGAGCCCATCTACAATAGCGAGGGGAAGCGGCTTAACACCCGAGAGTTCCGCACCC
          250      260      270      280      290      300

          70      80      90      100      110      120
TSAP2    GCAAAAAAAAAAATCTCTTGTTTTCTTAAGCTTTTCCCTGTGCTAGGGAAAGATCAGT
          :::::
humzfm1c.seq GCAAAAAGCTGGAAGAGGAGCGGCACAACCTCATCAGAGATGGTTGCACCTCAATCCGG
          310      320      330      340      350      360

          130      140
TSAP2    AAGTCCGTGGTTATAGATTGGTT

humzfm1c.seq ATTTCAAGCCACCTGCAGATTACAACCTCCAGCAACAGTGTGAGTGTAT
          370      380      390      400      410

```

FIG. 5

TSAP3

10

TSAP3 3

TTTTTTTTTTT

:::

mmsiah1b.seq TTGTAATAATTTCTGAACTTTGTATTTGTTGTAGATTGATTGTATTGTTGACAATTTT
1450 1460 1470 1480 1490 1500

20 30 40 50 60 70

TSAP 3 CCGGGTGGGGTGTGCCTGCACACATGCGTGCACGTGTGTGCTTGCTTTTCTTTAACAA
:::

mmsiah1b.seq CCGGGTGGGGTGTGCCTGCACACATGCGTGCACGTGTGTGCTTGCTTTTCTTTAACAA
1510 1520 1530 1540 1550 1560

80 90 100 110 120 130

TSAP 3 GCCATCTACGTGTGCATAGCCCACTGTTTTCCCTTGTGAGTCAACACATACTGCTGCTGT
:::

mmsiah1b.seq GCCATCTACGTGTGCATAGCCCACTGTTTTCCCTTGTGAGTCAACACATACTGCTGCTGT
1570 1580 1590 1600 1610 1620

140

TSAP3 CGTTTGGGTTTGGT

:::

mmsiah1b.seq CGTTTGGGTTTGGTTTGCTTTTGGTTTTTGATGTGTGTGATTTGATAATTTTATTCTA
1630 1640 1650 1660 1670 1680

FIG. 6

```

HUMSIAH      -----MSRQTATALPTGTSKCPPSQRVPAALTGTTASNN-----
MMSIAH1A_1   -----MSRQTATALPTGTSKCPPSQRVPAALTGTTASNN-----
MMSIAH1B_1   -----MSRQAATALSTGTSKCPPSQRVPAALTGTTASNN-----
DROSINA_1     MSNKIMPXRREPTAAAAGAGATGATNTSTSTGSSSSACNTSSANTSSSSSSSSLSAGGOD
               * * * * *

HUMSIAH      -----DLASLFECPVCFDYVLPPILOQCQSGHLVCSNCRPKLTCCPTCRGPLGSIKRLANE
MMSIAH1A_1   -----DLASLFECPVCFDYVLPPILOQCQSGHLVCSNCRPKLTCCPTCRGPLGSIKRLANE
MMSIAH1B_1   -----DLASLFECPVCFDYVLPPILOQCQSGHLVCSNCRPKLTCCPTCRGPLGSIKRLANE
DROSINA_1     AGMSADLTSLFECPVCFDYVLPPILOQCSSGHLVCSNCRPKLTCCPTCRGPLANIRLANE
               * * * * *

HUMSIAH      KVANSVLFPCKYASSGCEITLPHTKADHEELCEFRPYSCPCPGASCKWQGS L DAVM PHL
MMSIAH1A_1   KVANSVLFPCKYASSGCEITLPHTKAEHEELCEFRPYSCPCPGASCKWQGS L DAVM PHL
MMSIAH1B_1   KVANSVLFPCKYASSGCEITLPHTKAEHEELCEFRPYSCPCPGASCKWQGS L DAVM PHL
DROSINA_1     KVASNVKFPCKHSGYGCTASLVYTEXTEHEETCECRPYLCPCPGASCKWQGPLDLVMQHL
               * * * * *

HUMSIAH      MHQHKSIITTLQGEDIVFLATDINLP GAVDWVWQSCFGFHFMLVLEKQEKYDGHQOFFAI
MMSIAH1A_1   MHQHKSIITTLQGEDIVFLATDINLP GAVDWVWQSCFGFHFMLVLEKQEKYDGHQOFFAI
MMSIAH1B_1   MHQHKSIITTLQGEDIVFLATDINLP GAVDWVWQSCFGFHFMLVLEKQEKYDGHQOFFAI
DROSINA_1     MMSHKSIITTLQGEDIVFLATDINLP GAVDWVWQSCFGFHFMLVLEKQEKYDGHQOFFAI
               * * * * *

HUMSIAH      VOLIGTRKQAENFAYRLELNGHRRRLTWEATPRSIHEGIATAINMSDCLVFDT SIAQLFA
MMSIAH1A_1   VOLIGTRKQAENFAYRLELNGHRRRLTWEATPRSIHEGIATAINMSDCLVFDT SIAQLFA
MMSIAH1B_1   VOLIGTRKQAENFAYRLELNGHRRRLTWEATPRSIHEGIATAINMSDCLVFDT SIAQLFA
DROSINA_1     VOLIGSRKEAENFVYRLELNGNRRRLTWEAMP RSIHEGVASATHNSDCLVFDT SIAQLFA
               * * * * *

HUMSIAH      ENGNLGINVTISM C
MMSIAH1A_1   ENGNLGINVTISM C
MMSIAH1B_1   ENGNLGINVTISM C
DROSINA_1     DNGNLGINVTISLV
               * * * * *

```

FIG. 7

1 mms182	10	20	30
2 tsip2	40	50	60
1 mms182	70	80	90
2 tsip2	100	110	120
1 mms182	130	140	150
2 tsip2	160	170	180
1 mms182	190	200	210
2 tsip2	220	230	240
1 mms182	250	260	270
2 tsip2	280	290	300
1 mms182	310	320	330
2 tsip2			

FIG. 8

1 mms182	340	350	360
2 tsip2	tcctggagcca gaatgacagc caagaacggc		
	TCCGGAGCCA GAATGACAGC CAAGAACGGC		
1 mms182	370	380	390
2 tsip2	agcagcagca tgacagggcag agacttcgaca		
	AGCAGCAGCA TGACAGGCAG AGACTTCACA		
1 mms182	400	410	420
2 tsip2	accctgagcc aatatctaac gggcggtcccc		
	ACCCTGAGCC AATATCTAAT GGGCGGCCCC		
1 mms182	430	440	450
2 tsip2	agagtaactc aagacaggtg gcggaacaag		
	AGAGTAACTC AAGACAGGTG GTGGAACAAG		
1 mms182	460	470	480
2 tsip2	atgaggagga agacyaagag ctgacattga		
	ATGAGGAGGA AGACGAAGAG CTGACATTGA		
1 mms182	490	500	510
2 tsip2	aatatggagc caagcatgct atcatgctct		
	AATATGGAGC CAAGCATGCT ATCATGCTCT		
1 mms182	520	530	540
2 tsip2	ctgtccccgt gacccctctgc atggctcgtag		
	TTGTCCCCGT GACCCCTCTGC ATGGCTCGTAG		
1 mms182	550	560	570
2 tsip2	tcggggccac catcaaatca gtcagcttct		
	TCCGTGCCAC CATCAAAATCA GTCAGCTTCT		
1 mms182	580	590	600
2 tsip2	ataccgggaa ggacgggtcag ccaatctaca		
	ATACCCGGAA GGACGGGTCA CTAATCTACA		
1 mms182	610	620	630
2 tsip2	ccccattcac ageagacact gagactgtag		
	CCCCATTCA CAGAGACACT GAGACTGTAG		
1 mms182	640	650	660
2 tsip2	gccccagaga cctgcaactcg atccctgaatg		
	GCCAAAGAGC CCTGCACTCG ATCCTGAATG		

FIG. 8 (suite)

FEUILLE DE REMPLACEMENT (REGLE 26)

1 mms182	670 680 690	cgcccatcsc gaccagcgcc attgtcatta
2 tsip2		CGGCCATCAT GATCAGTGTG ATTGTGATTA
1 mms182	700 710 720	tgaccatcsc cctggcgcc cgtataaac
2 tsip2		TGACCATCCT CCTGGTGTG CTGTATAAAT
1 mms182	730 740 750	acaggcgcc caaggcgcc cagcgctggc
2 tsip2		ACAGGTGCTA CAAGGTGATC CAGCGCTGCC
1 mms182	760 770 780	ccattatccc atctcgtcg cgcggttcc
2 tsip2		TTATTATTTC ATCTCTGTTG TTGCTGTTCT
1 mms182	790 800 810	ccccctcgtc cacttactca ggggaagcat
2 tsip2		TTTTTTCGTT CATTTACTTA GGGGAAGTAT
1 mms182	820 830 840	ccaagacctt caatgcggcc ggggaactacg
2 tsip2		TTAAGACCTA CAATGTGCGC GTGGACTAAG
1 mms182	850 860 870	ccacagtagc acccttaatc tgggaattctg
2 tsip2		TTACAGTAGC ACTCTTAATC TGGGAATTTC
1 mms182	880 890 900	gtgtggcgcc gctgactgac atccactgga
2 tsip2		GTGTGGTGGC GATGATTGCC ATCCACTGGA
1 mms182	910 920 930	aaggcccccc cggactgac caggcgatc
2 tsip2		AAGGCCCCCT TCGACTGAC CAGCGGTATC
1 mms182	940 950 960	ccattatgat cagtgccttc atggccctgg
2 tsip2		TCATTATGAT CAGTGCCTTC ATGCCCTGCG
1 mms182	970 980 990	cattcatcaa gtaacctccc gaatggaccg
2 tsip2		TATTTATCAA GTACCTCCCC GAATGGACCC

FIG. 8 (suite)

1 mms182	1000	1010	1020
2 tsip2	catggctcct	cttggtcttg	attccagcat
1 mms182	1030	1040	1050
2 tsip2	atgattctgg	ggctgcttta	tgccccaaag
1 mms182	1060	1070	1080
2 tsip2	gccccacttcg	catgctggcc	gaaacagctc
1 mms182	1090	1100	1110
2 tsip2	aggaagaaaa	cgagactctc	ctccagctc
1 mms182	1120	1130	1140
2 tsip2	ttatctattc	ctcaacaatg	gtgtggttgg
1 mms182	1150	1160	1170
2 tsip2	tgaatctggc	tgaaggagac	ccagaagucc
1 mms182	1180	1190	1200
2 tsip2	aaaggagggc	acccaagaaac	cccaagtata
1 mms182	1210	1220	1230
2 tsip2	acacacaaaag	agcggagaga	gagacacagg
1 mms182	1240	1250	1260
2 tsip2	acagtggttc	tgggaacgat	gacggctggc
1 mms182	1270	1280	1290
2 tsip2	tcagtgaaga	gtgggagggc	caaagagaca
1 mms182	1300	1310	1320
2 tsip2	gtccacctggg	gcttcctcgg	tccactcccg
2 tsip2	gtcacttggg	gcttcctcgg	tccactcccg

FIG. 8 (suite)

1 mms182	1330	1340	1350
2 tsip2	agccaagagc	tgctgcccag	gaacccccctg
1 mms182	1360	1370	1380
2 tsip2	AGTCAAGAGC	TGCTGTCCAG	GAACCTTCTG
1 mms182	1390	1400	1410
2 tsip2	ggagcattct	aacgagtga	gacccggagg
1 mms182	1420	1430	1440
2 tsip2	GGAGCATTCT	AACGAGTGAA	GACCCGGAGG
1 mms182	1450	1460	1470
2 tsip2	aaagaggagt	aaaacclgga	ctgggagatt
1 mms182	1480	1490	1500
2 tsip2	AAAGAGGAGT	AAAACCTTGA	CTGGGAGATT
1 mms182	1510	1520	1530
2 tsip2	tcattttcta	cagugccctg	gctggtaagg
1 mms182	1540	1550	1560
2 tsip2	TCATTTTCTA	CAGTGTCTG	CTTGTAAGG
1 mms182	1570	1580	1590
2 tsip2	cctcagcaac	ggcaglyga	gactggaaca
1 mms182	1600	1610	1620
2 tsip2	CCTCAGCAAC	GCCAGTGA	GACTGGAACA
1 mms182	1630	1640	1650
2 tsip2	caaccatagc	ctgccttgta	gccataactga
1 mms182	1660	1670	1680
2 tsip2	CAACCATAGC	CTGCTTTGTA	GCCATACTGA
1 mms182	1690	1700	1710
2 tsip2	tcggccctgtg	ccttacatta	ctcctgctcg
1 mms182	1720	1730	1740
2 tsip2	TGGGCCCTGTG	CCTTACATTA	CTCCTGCTCG
1 mms182	1750	1760	1770
2 tsip2	ccattttcaa	gaaagcgctg	ccagccctcc
1 mms182	1780	1790	1800
2 tsip2	CCATTTTCAA	GAAAGCGCTG	CCAGCCCTCC
1 mms182	1810	1820	1830
2 tsip2	ccatctccat	caucttccgg	ctcctgttct
1 mms182	1840	1850	1860
2 tsip2	CCATCTCCAT	CAUCTTCCGG	CTCCTGTTCT
1 mms182	1870	1880	1890
2 tsip2	acttcgccac	ggattacccc	gtgcagcccc
1 mms182	1900	1910	1920
2 tsip2	ACTTCGCCAC	GUATTACCTT	GTGCAGCCCT
1 mms182	1930	1940	1950
2 tsip2	tcattggacca	acttgcattc	cattcagcttc
1 mms182	1960	1970	1980
2 tsip2	TCATTGGACCA	ACTTGCATTG	CATCAGTTT

FIG. 8 (suite)

1 mms182	1650	1670	1680
2 tsip2	atattccagcc	cttcctgcagc	tagaacacgg
	ATATCTAGCC	TTCCTGCCAGT	TAGAACATGG
1 mms182	1690	1700	1710
2 tsip2	atgtttcttc	tttgattatc	aaaaacacaa
	ATGTTTCTTC	TTTGATTATC	AAAAACACAA
1 mms182	1720	1730	1740
2 tsip2	aaacagagag	caagcccgag	gaggagactg
	AAACAGAGAG	CAAGCCCGAG	GAGGAGACTG
1 mms182	1750	1760	1770
2 tsip2	gtgactttcc	ttgtgcttca	gctaacaag
	GTGACTTTCC	TGTGCTCTCA	GCTAACAAAG
1 mms182	1780	1790	1800
2 tsip2	gcaggacttc	agctggactc	ctgcagcttc
	GCAGGACTTC	AGCTGGACTT	CTGCAGCTTC
1 mms182	1810	1820	1830
2 tsip2	cttcctagacc	cttcctagcca	cttcctagctac
	CTTCCTAGACC	CTTCCTAGCCA	CTTCCTAGCTAC
1 mms182	1840	1850	1860
2 tsip2	ctgactctgg	aagggaagcg	ctacagagga
	CTGACTCTGG	AAGGAAGCGT	CTACAGAGGA
1 mms182	1870	1880	1890
2 tsip2	acggctctcc	acatccatcg	ctgcagcaga
	ACGCTTTCCA	ACATCCATCG	CTGCAGCAGA
1 mms182	1900	1910	1920
2 tsip2	cggctgctcc	cagtgacttg	agagacaagg
	CGGTGCTCCT	CAGTGACTTG	AGAGACAAGG
1 mms182	1930	1940	1950
2 tsip2	acaaggaaac	gtgctggggc	aaggagctgc
	ACAAGGAAAT	GTCCTGGGCC	AAGGAGCTGC
1 mms182	1960	1970	1980
2 tsip2	cgtgctctcc	tagctctggg	cgtggggcag
	CGTGCTCTCC	TAGCTCTGGG	CGTGGGGCAG

FIG. 8 (suite)

1 mms182	1990	2000	2010
2 tsip2	gagattccacc cggactctga cccccctaaag GAGATTTACC CGCACTTGTGA ACTCTCTAAG		
1 mms182	2020	2030	2040
2 tsip2	gtaaacaaaag cggagctgac c GTAAACAAAAG TGAGCTGAAC CAAACAGAGC		
1 mms182	2050	2060	2070
2 tsip2	<== <== TGCCATYCTT CCACACCATG TTGGAAATAA		
1 mms182	2080	2090	2100
2 tsip2	<== <== AACCGTCCTA GCTGGAACCC TTA CTGTCCC		
1 mms182	2110	2120	2130
2 tsip2	<== <== AGGAGGTTCC GTGTGGGGGT GGCACCTGGGC		
1 mms182	2140	2150	2160
2 tsip2	<== <== CGGGCCTCCC TCTCAGGCTC CTTTGCTGCC		
1 mms182	2170	2180	2190
2 tsip2	<== <== CACTTGTAAG TTAAATAAG GACACCCCCC		
1 mms182	2200	2210	2220
2 tsip2	<== <== TACACAAACC TCACCCCTGT CACATCCAGT		
1 mms182	2230	2240	2250
2 tsip2	<== <== GACTCTGACC ACTTTAGTTC TCAAACTCTC		
1 mms182	2260	2270	2280
2 tsip2	<== <== TCACTATTAT CAGTGCTTCC CTTTTCTTCC		
1 mms182	2290	2300	2310
2 tsip2	<== <== CAAGGCCAAG CTGCACCAAT TTGGCGTTCC		

FIG. 8 (suite)

1 mms182	2320	2330	2340
3	<==		
2 tsip2	<==		
	TCTATCCTGA GAATTGTAAC CTCAACTTCC		
1 mms182	2350	2360	2370
3	<==		
2 tsip2	<==		
	AAAGTTTATA TTTTCCTGAA ATGATGGATC		
1 mms182	2380	2390	2400
3	<==		
2 tsip2	<==		
	TATTGCTCAA CAGTCCCTGT CATCCTTAAG		
1 mms182	2410	2420	2430
3	<==		
2 tsip2	<==		
	TGACTTCTGG GTTCCCAACA AATTCCTCAC		
1 mms182	2440	2450	2460
3	<==		
2 tsip2	<==		
	TTTLAGACAC ACTCTAAGCT TACTTCTGGC		
1 mms182	2470	2480	2490
3	<==		
2 tsip2	<==		
	CTGGATGCTT CCTCTCCCTG TCTCTCCCTT		
1 mms182	2500	2510	2520
3	<==		
2 tsip2	<==		
	GCCCCACAGC GGTTCCTGA CAGCAGACAA		
1 mms182	2530	2540	2550
3	<==		
2 tsip2	<==		
	GGCAGCTCTG GGAGGTAGCT AGTATCCAAT		
1 mms182	2560	2570	2580
3	<==		
2 tsip2	<==		
	AACCCAGGGG TTTCTCATG TGATGCAAAT		
1 mms182	2590	2600	2610
3	<==		
2 tsip2	<==		
	ACTACGTGTG CAACCAATCA GTGCTGTCAA		
1 mms182	2620	2630	2640
3	<==		
2 tsip2	<==		
	CGGGCTGCCA TAGCTCCTTC GATGGCAAAT		

FIG. 8 (suite)

		2550	2600	2650
1 mms182		<==		
3		<==		
2 tsip2		AGGATGTGTC CCCAAAGAAT TAAAGCGATC		
		2630	2690	2700
1 mms182		<==		
3		<==		
2 tsip2		ACTGGCTGGT G		

FIG. 8 (fin)